



Journal of
*Risk and Financial
Management*

Special Issue Reprint

Commodity Market Finance

Edited by
Kentarō Iwatsubo

mdpi.com/journal/jrfm



Commodity Market Finance

Commodity Market Finance

Editor

Kentaro Iwatsubo



Basel • Beijing • Wuhan • Barcelona • Belgrade • Novi Sad • Cluj • Manchester

Editor

Kentaro Iwatsubo
Graduate School of Economics
Kobe University
Kobe
Japan

Editorial Office

MDPI
St. Alban-Anlage 66
4052 Basel, Switzerland

This is a reprint of articles from the Special Issue published online in the open access journal *Journal of Risk and Financial Management* (ISSN 1911-8074) (available at: www.mdpi.com/journal/jrfm/special-issues/Commodity_Market_Finance).

For citation purposes, cite each article independently as indicated on the article page online and as indicated below:

Lastname, Firstname, Firstname Lastname, and Firstname Lastname. Article Title. <i>Journal Name</i> Year , <i>Volume Number</i> , Page Range.
--

ISBN 978-3-0365-9029-5 (Hbk)

ISBN 978-3-0365-9028-8 (PDF)

doi.org/10.3390/books978-3-0365-9028-8

© 2023 by the authors. Articles in this book are Open Access and distributed under the Creative Commons Attribution (CC BY) license. The book as a whole is distributed by MDPI under the terms and conditions of the Creative Commons Attribution-NonCommercial-NoDerivs (CC BY-NC-ND) license.

Contents

About the Editor	vii
Preface	ix
Hiroyuki Okawa	
Markov-Regime Switches in Oil Markets: The Fear Factor Dynamics Reprinted from: <i>J. Risk Financial Manag.</i> 2023 , 16, 67, doi:10.3390/jrfm16020067	1
Kentaro Iwatsubo and Clinton Watkins	
Causality between Arbitrage and Liquidity in Platinum Futures Reprinted from: <i>J. Risk Financial Manag.</i> 2022 , 15, 593, doi:10.3390/jrfm15120593	21
Faheem Aslam, Paulo Ferreira and Haider Ali	
Analysis of the Impact of COVID-19 Pandemic on the Intraday Efficiency of Agricultural Futures Markets Reprinted from: <i>J. Risk Financial Manag.</i> 2022 , 15, 607, doi:10.3390/jrfm15120607	38
Md. Kausar Alam, Mosab I. Tabash, Mabruk Billah, Sanjeev Kumar and Suhaib Anagreh	
The Impacts of the Russia–Ukraine Invasion on Global Markets and Commodities: A Dynamic Connectedness among G7 and BRIC Markets Reprinted from: <i>J. Risk Financial Manag.</i> 2022 , 15, 352, doi:10.3390/jrfm15080352	56
Hector O. Zapata, Junior E. Betanco, Maria Bampasidou and Michael Deliberto	
A Cyclical Phenomenon among Stock & Commodity Markets Reprinted from: <i>J. Risk Financial Manag.</i> 2023 , 16, 320, doi:10.3390/jrfm16070320	76
Afees A. Salisu, Rangan Gupta and Riza Demirer	
Oil Price Uncertainty Shocks and Global Equity Markets: Evidence from a GVAR Model Reprinted from: <i>J. Risk Financial Manag.</i> 2022 , 15, 355, doi:10.3390/jrfm15080355	89
Nobuhiro Nakamura, Kazuhiko Ohashi and Daisuke Yokouchi	
Dynamic Relationship between Volatility Risk Premia of Stock and Oil Returns Reprinted from: <i>J. Risk Financial Manag.</i> 2023 , 16, 173, doi:10.3390/jrfm16030173	115
Mehmet Balcilar and Huseyin Ozdemir	
On the Risk Spillover from Bitcoin to Altcoins: The Fear of Missing Out and Pump-and-Dump Scheme Effects Reprinted from: <i>J. Risk Financial Manag.</i> 2023 , 16, 41, doi:10.3390/jrfm16010041	135
Junwei Chen	
Analysis of Bitcoin Price Prediction Using Machine Learning Reprinted from: <i>J. Risk Financial Manag.</i> 2023 , 16, 51, doi:10.3390/jrfm16010051	150
Jin Shang and Shigeyuki Hamori	
Do Large Datasets or Hybrid Integrated Models Outperform Simple Ones in Predicting Commodity Prices and Foreign Exchange Rates? Reprinted from: <i>J. Risk Financial Manag.</i> 2023 , 16, 298, doi:10.3390/jrfm16060298	175
Kentaro Hoshisashi and Yuji Yamada	
Pricing Multi-Asset Bermudan Commodity Options with Stochastic Volatility Using Neural Networks Reprinted from: <i>J. Risk Financial Manag.</i> 2023 , 16, 192, doi:10.3390/jrfm16030192	200

About the Editor

Kentaro Iwatsubo

Kentaro Iwatsubo was born in Nishinomiya City, Hyogo Prefecture, in 1969. He graduated from the School of Political Science and Economics at Waseda University in 1993, received a Master's degree from the Graduate School of Economics at the University of Tokyo in 1997, and received a Ph.D. in Economics from the University of California, Los Angeles (UCLA) in 2003. After working as an Assistant Professor at the Institute of Economic Research, Hitotsubashi University, he became an Associate Professor at the Graduate School of Economics, Kobe University, in 2007 and a Professor in 2013. During that time, he was a visiting professor at the University of Michigan, USA; Sung Kyun Kwan University, Korea; and Cass Business School, UK. For many years, he was a Special Research Fellow at the Ministry of Finance's Policy Research Institute. He has served as an executive board member of the Japan Society of Monetary Economics and as Editor-in-Chief of the *Japanese Journal of Monetary and Financial Economics*. He is currently a board member of the Japan Finance Association and an academic advisor to the Financial Futures Association of Japan. He has published many research articles in the areas of foreign exchange, stocks, and bonds, as well as commodities. He received the Academic Encouragement Award from Murao Ikueikai in 2015 for his analysis of exchange rate finance.

Preface

The world of commodity markets has undergone a major transformation over the past few decades, taking a major turn toward financialization. Since the early 2000s, commodities, once the domain of farmers, miners, and physical traders, have become increasingly important to investors, speculators, and financial institutions. This evolution has created a dynamic and ever-changing arena in which the various factors affecting the market are intricately intertwined.

The dawn of the 21st century brought a series of events that had a profound impact on commodity markets. Rising international tensions, exemplified by the Cold War between the U.S. and China, highlighted the geopolitical nature of commodity markets. Meanwhile, the sudden and devastating outbreak of COVID-19 shocked the global economy, disrupting supply chains and causing roller coaster-like price volatility. In addition, Russia invaded Ukraine, and uncertainty spilled over into the markets, sending resource prices soaring.

These cataclysmic events amplified the volatility inherent in commodity markets. At the same time, financial markets such as equities, bonds, and foreign exchange also showed volatility. The interaction between commodity and financial markets deepened and formed a complex nexus that came under close scrutiny from governments, corporations, and investors around the world.

This reprint delves into recent developments in the commodity markets and elucidates the multifaceted factors that have shaped their trajectory. It examines how the interwoven dynamics of supply and demand, geopolitics, technology, and financialization have brought about a new era in commodity trading. By providing a comprehensive survey of these developments, we aim to provide insights that will help stakeholders successfully navigate the challenges and opportunities presented by this evolving landscape.

We invite you to join us on this journey as we explore recent transformations in commodity markets and their far-reaching implications for the global economy. From the rise of renewable energy to the impact of climate change on commodity prices, we will traverse the complex web of factors that define this dynamic and evolving field.


Special thanks are in store for all the authors involved in the publication of this reprint. I am also deeply grateful to Ms. Chelthy Cheng, Managing Editor of MDPI, for her help in planning this Special Issue. Finally, I would like to thank my wife Junko and my son Takumi for their support in my research life.

Kentaro Iwatsubo

Editor

Article

Markov-Regime Switches in Oil Markets: The Fear Factor Dynamics

Hiroyuki Okawa 

Graduate School of Economics, Kobe University, Kobe 657-8501, Japan; phiro.lpworld.bj@gmail.com

Abstract: This paper is an attempt to examine regime switches in the empirical relation between return dynamics and implied volatility in energy markets. The time-varying properties of the return-generating process are defined as a function of several risk factors, including oil market volatility and changes in stock prices and currency rates. The empirical evidence is based on Markov-regime switching models, which have the capacity to capture, in particular, the stochastic behavior of the OVX oil volatility index as a benchmark for investors' fear. The results suggest that the dynamics of oil market returns are governed by two distinct regimes, a state driven by a negative relationship between returns and implied volatility and another state characterized by a more pronounced negative correlation. It is the latter regime with a stronger correlation that tends to prevail over the sample period from 2008 to 2021, but the frequency of regime shifts also seems to increase under more volatile oil price dynamics in association with significant events such as the COVID-19 pandemic. Thus, the evidence of a negative correlation structure is found to be robust to changes in the estimation period, which suggests that the oil volatility index remains a reliable gauge of market sentiment in the energy markets.

Keywords: energy market volatility; oil price dynamics; fear index; Markov-regime switching models



Citation: Okawa, Hiroyuki. 2023. Markov-Regime Switches in Oil Markets: The Fear Factor Dynamics. *Journal of Risk and Financial Management* 16: 67. <https://doi.org/10.3390/jrfm16020067>

Academic Editor: Thanasis Stengos

Received: 26 December 2022

Revised: 20 January 2023

Accepted: 20 January 2023

Published: 23 January 2023



Copyright: © 2023 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The dynamics of energy markets have a strong bearing not only on various aspects of social life but also on economic activities, monetary policies, and investment decisions. Price signals from the crude oil market, in particular, have significant effects on the behavior of inflation expectations and, in turn, institutional investors and policymakers. It may also be argued that the price dynamics are reflective of the aggregate impact of megatrends, including demographic and social changes, technological innovation, natural resources, financial globalization, rapid urbanization, and shifts in economic power, inter alia. It is not clear how this complex web of underlying forces may affect the energy markets and, for instance, their essential linkage with the equity and currency markets. The existence of latent variables has the potential to create non-linear relationships that may not be easily reflected by linear regression models. Thus, shifts in the relationship between price variations in energy markets and equity returns, as well as currency changes between different states of the latent variables, may be better captured by Markov-regime switching models. It is the principal objective of the present study to examine the inter-market linkages as well the inner dynamics of the risk-return tradeoff relationship in oil futures markets.

Conventional wisdom suggests that market perceptions of increased economic uncertainty are likely to be accompanied by lower asset valuation and expectations of higher volatility and increasing volatility. The negative relation between returns and changes in volatility expectations in equity markets is usually assessed using model-free volatility benchmarks such as the Chicago Board Options Exchange's VIX index, which is regarded as a gauge of investors' fear. The fear factor dynamics, measured by changes in the OVX volatility index, may constitute a significant determinant of shifts in the risk-return trade-off in oil futures markets. The OVX volatility index is derived from the option prices

of Exchange Traded Funds that are linked to spot WTI prices and provides a forward-looking measure of market volatility in terms of future oil price fluctuations. The focus on the risk-return dynamics of WTI markets is justified by the crucial role that energy markets play in the real economy and their intrinsic relationship with financial markets. Indeed, unlike financial markets, commodity markets, including crude oil futures, have a tendency to exhibit price seasonality, reflecting shifting risks associated with oil production and consumption due to changes in supply and demand functions, in addition to geopolitical risks.

Thus, a better understanding of the dynamics of oil futures returns may shed light on the significance of market sentiment and investors' fear on the linkage between the real economy and financial economy, and particularly on the impact of the compounded healthcare and economic crises. Indeed, the COVID-19 pandemic was characterized by expectations of heightened economic uncertainty and sharp falls in WTI futures, with negative pricing reflecting the effects of economic lockdowns in terms of the underutilization of production factors and disruptions to supply chains. The econometric approach is based on Markov-regime switching models, which have the capacity to capture abrupt changes in the correlation structure and the propensity of oil markets as well as financial markets to behave differently during periods of lower and higher economic uncertainty. Thus, Markov-regime switching models can indeed be useful in better understanding the shifting behavior of energy markets and anticipating changes in the correlation structure in response to significant events.

To the best of the author's knowledge, this paper provides new evidence about the prevalence of a Markov regime of stronger negative correlation under more volatile markets. The empirical results suggest that futures returns are governed by two latent states exhibiting significant negative correlations between WTI futures returns and OVX daily changes. The new evidence indicates that regime shifts are more likely to occur in association with market expectations of increasing volatility and diminishing oil returns. Futures returns during periods of increased uncertainty tend to be governed by abrupt switches between regimes of weaker and stronger negative correlations with OVX, reflecting changing levels of investors' fear. In addition, the new evidence suggests that economic lockdowns in response to disease outbreaks have the potential to increase the likelihood of Markov regimes with a more pronounced negative correlation between oil futures returns and changes in volatility expectations. In contrast, periods of financial instability, such as the U.S. credit crisis, have the potential to weaken or impair the inherent relationship of oil futures returns with volatility expectations and instead strengthen the linkages with currency fluctuations and equity valuation.

The remainder of the paper is organized as follows. Section 2 provides a brief review of the literature related to the VIX model-free volatility index and the empirical evidence about its correlation to equity returns. Another strand of the literature is related to the OVX volatility-return inner dynamics of the WTI futures market, as well as its correlation with other asset markets. Section 3 describes the Markov-Regime Switching modeling of WTI return dynamics. Section 4 presents the sample data, including WTI futures, the oil volatility index, S&P 500 stock price index, and the U.S. dollar index. It describes the distributional properties and discusses the estimation results of Markov-regime switching models for the full sample period. The empirical evidence is also inclusive of robustness tests required to examine the stability of the correlation structure over time. The final section concludes the paper.

2. Literature Review

The model-free methodology underlying the calculation of the VIX volatility index by the Chicago Board Options Exchange (CBOE) is shared by many other volatility indices, including the OVX index for crude oil futures markets. The volatility index provides a forward-looking measure of market sentiment in terms of expectations about future levels of price volatility. The VIX index reflects the volatility implicit in a hypothetical

option on the S&P 500 index, assuming a constant time to expiration of thirty days. This volatility index is also widely regarded as a measure of investors' fear because of its negative correlation with returns. There is a strong tendency for the anticipated level of volatility to increase under bearish markets and decrease under bullish markets. Theoretically, the sensitivity of volatility expectations to fluctuations in the price of the underlying asset is an intrinsic feature of option pricing. Thus, perceptions of increasing economic uncertainty by options market participation are bound to be accompanied by lower asset valuations and expectations of higher volatility in the underlying asset market. This proposition about the return-volatility dynamics is expected to apply with equal force in options and underlying markets, independent of the nature of assets, including equities and commodities.

Earlier studies by Fleming et al. (1995), and Connolly et al. (2005), among others, provided evidence of a strong negative correlation between the VIX index and stock market returns.¹ Sarwar (2012) presented further evidence of this empirical relationship based on contemporaneous variables and suggested that a decline in equity prices is conducive to market perceptions of increased uncertainty and, in turn, higher volatility expectations. The original work by Whaley (2000, 2009), and more recent studies by Smales (2022), among others, indicate that the VIX functions as a fear index for market investors. The development of comparable volatility indices for other markets, such as the VXJ index from the Nikkei 225 option prices by Nishina et al. (2006) for the Japanese markets and from the Kospi 200 options for the Korean markets by Maghrebi et al. (2007) provided additional evidence about the usefulness of the model-free implied volatility index as a gauge of investors' fear and market sentiment. An alternative version of model-free volatility is proposed by Fukasawa et al. (2011) based on the approximation of the expected quadratic variations of asset prices in relation to options prices.²

Part of the empirical literature also focuses on the stochastic behavior of volatility indices in relation to asset bubbles, financial crises, and macroeconomic shocks. For instance, some empirical evidence is provided by Szado (2009), Nishina et al. (2012), Maghrebi et al. (2014), and Baiardi et al. (2020), among others, with respect to credit crises and financial instability. Earlier evidence from Giot (2003) and Maghrebi et al. (2007) sheds light on regime switches in relation to the Asian currency crisis. More recent studies by Just and Echaust (2020) and Grima et al. (2021) provide evidence about the behavior of volatility expectations in association with the COVID-19 disease outbreak. Thus, many empirical studies of the non-linear dynamics of volatility expectations in equity options markets are based on Hamilton's modeling of Markov-regime switches.

Another strand of the literature focuses on the relevance of commodity markets, including crude oil, in explaining the behavior of returns in equity, bond, and foreign exchange markets.³ Part of the reason for the focus of empirical analysis on energy markets lies in the need to examine the effectiveness of including commodity markets for portfolio risk diversification purposes. In this context, the study by Beckmann et al. (2020) examines the relationship between WTI futures and exchange rates which reflects the extent of international trade. Given their importance for the real economy and relationship with financial markets, crude oil markets are the subject of a growing literature aimed at better understanding the price and volatility dynamics. The OVX volatility benchmark is an index calculated based on the CBOE's VIX methodology, but it is rather derived from the option prices of Exchange Traded Funds (ETFs) that are intrinsically linked to WTI option prices. Thus, in a sense, the OVX index is an aggregation of volatility expectations by participants in the ETFs rather than crude oil markets.

Crude oil markets are intrinsically different from equity markets. Commodity prices exhibit seasonality and are strongly sensitive to economic activities, with risks associated with the production and consumption functions. Crude oil is traded as a real asset, and market prices are sensitive to a delicate balance between supply and demand, which is influenced by geopolitical risks, among others. The literature includes many studies, including Pindyck (2001), Hamilton (2008), and Gong and Xu (2022), among others, that consider the dynamics and determinants of commodity markets as well as the impact of

geopolitical risk. It is also noted that market participants tend to trade crude oil not just to facilitate economic activities but for speculative purposes as well. Thus, it is important to understand the nature of crude oil markets and their stochastic behavior under different levels of economic uncertainty.

There is a rich body of literature on the nature of commodity exchanges and their relationship with financial markets. The empirical evidence from Dupoyet and Shank (2018) suggests that changes in the OVX index are positively related to stock market returns in the manufacturing, energy, and utilities sectors and negatively related to those in durable consumer goods and wholesale trade sectors. Earlier evidence from Bodie and Rosansky (1980) suggests that it is possible to make recourse to alternative investments to hedge risk in financial markets by focusing on the nature of commodities. More recent studies by Tang and Xiong (2012), Silvennoinen and Thorp (2013), and Cheng and Xiong (2014) provide evidence about a process of financialization of commodity markets reflecting the prevailing structures of stock markets. The impact of financialization on commodity markets is shown by Goldstein and Yang (2022) to be linked with the real economy. It is important to understand also the impact of energy price fluctuations on the real economy and the behavior of financial markets and institutions. As noted in Dutta (2017), the observed levels of oil market volatility are significantly higher than price fluctuations in stock markets. Additionally, the Federal Reserve Board (2022) argues that the volatility in commodity markets, including energy resources and food, may pose significant risks to the stability of the financial system. A heightened systemic risk is reflective of the sensitivity of the real economy to unexpected variations in energy and commodity markets.

Other earlier studies by Bodie and Rosansky (1980), Cheung and Miu (2010), and Jensen et al. (2000), among others, examined the effectiveness of risk diversification based on commodities markets. Additionally, Gorton and Rouwenhorst (2006) provide evidence that the returns on commodity futures tend to be negatively correlated with stock market returns and bond returns and that commodity futures are positively correlated with inflation as well as unanticipated inflation and changes in expected inflation. Given these stylized facts about commodity futures, further research has shed more light on the correlation structure between commodities and other asset markets over different time periods. For instance, Lombardi and Ravazzolo (2013) developed a Bayesian Dynamic Conditional Correlation model that can account for time-varying correlation patterns between commodity and equity returns and show that it is possible to obtain more accurate density forecasts. Furthermore, Baumeister and Kilian (2012, 2015) on oil price forecasting. There is also a growing field of literature using Machine Learning and deep learning in an attempt to obtain more accurate forecasts. The crucial importance of crude oil markets and their linkages with alternative markets is manifest in several empirical studies, including the work by Ferraro et al. (2015), who examine the potential to predict exchange rates from crude oil prices.

Accordingly, there is a growing body of literature on the return dynamics of WTI futures and their relationship with volatility expectations and other risk factors, including fluctuations in stock prices and exchange rates. For instance, Aboura and Chevallier (2013) found a positive correlation, or inverse leverage effects, between changes in OVX levels and oil prices in association with the onset of the U.S. credit crisis. The evidence suggests that the positive relationship may be reflective of consumers' fear of higher oil prices, which stands in sharp contrast with the nature of fear in equity markets about downside price movements. The empirical study by Chen et al. (2015) suggests, however, that WTI futures returns and OVX daily changes are negatively correlated over a sample period, partly overlapping with that of the study by Aboura and Chevallier (2013). Further empirical studies by Liu et al. (2017) and Dupoyet and Shank (2018) present similar evidence of negative correlation or asymmetric dependence based on mixed copula and GJR-GARCH models, respectively. The latter study also suggests that volatility expectations measured by the OVX index have a greater influence on financial markets than oil prices themselves and that both oil volatility and prices are significantly related to corporate bond credit spreads.

Thus, given the mixed evidence on the dynamics of risk-return tradeoff and the growing literature on the proposition that oil volatility expectations are also contingent on returns in foreign exchange and financial markets, it is important to explore the non-linear dynamics of the WTI futures returns and OVX index over more recent periods, including the long-term effects of financial crises and disease outbreaks. The present study uses Markov-regime switching models, which have the capacity to capture abrupt changes in the correlation structure and the propensity of oil markets, as well as financial markets, to behave differently during periods of lower and higher economic uncertainty. Of particular interest is the market behavior during the COVID-19 pandemic, with heightened volatility expectations and negative pricing of WTI futures, which are reflective of the effects of economic lockdowns in terms of the underutilization of production factors and disruptions to supply chains. Modeling the return dynamics of oil futures with Markov-regime switching models can be useful in better understanding the stochastic behavior of energy markets, which are different in nature from financial markets but may exhibit similar regime shifts in response to significant events.

3. Markov-Regime Switching Modeling of Return Dynamics in Oil Futures Markets

The empirical analysis of the return dynamics of crude oil futures is based on the Markov-regime switching model proposed by Hamilton (1989). Futures returns $y_{WTI,t}$ can be simply expressed with a first-order autoregression

$$y_{WTI,t} = \omega_{s_t} + \alpha_{s_t} y_{WTI,t-1} + \varepsilon_t \quad (1)$$

where the disturbance terms are white noise distributed with $\varepsilon_t \sim i.i.d.N(0, \sigma^2)$. The drift term ω_{s_t} and auto-regressive coefficient α_{s_t} are assumed to depend on the state s_t prevailing at time t . This allows for the intercept value, for instance, to change from ω_1 to ω_2 with an imperfectly predictable change in the average level of the return series $y_{WTI,t}$ from one state to another at time t . Similarly, the auto-regressive term α_{s_t} , which reflects the degree of mean reversion or long memory, is allowed to adapt to changes in the stochastic structure of the return series.

Following the empirical study by Sarwar (2012) focusing on the dynamics of the S&P 500 returns as a function of contemporaneous changes in the VIX index, respectively, it is possible to express the stochastic properties of oil futures returns $y_{WTI,t}$ in Equation (1) as a function of changes in the OVX index $y_{OVX,t}$ as well. Additionally, given the empirical evidence based on Markov-regime switching models estimated by Baiardi et al. (2020) that S&P 500 returns are likely to be negatively related to oil futures returns, it is important to account for the linkage between oil futures and equity prices. Finally, there is a need to also examine the impact of exchange rate fluctuations given the fact that WTI futures are denominated in U.S. dollars and in light of evidence from Beckmann et al. (2020) that oil futures are sensitive to currency fluctuations.

Thus, it is possible to describe the dynamics of oil futures returns $y_{WTI,t}$ as a function of the inner autoregressive terms, changes in the OVX index as a measure of investors' fear in oil markets $y_{OVX,t}$, as well as returns in currency and equity markets, expressed by U.S. dollar index returns $y_{USD,t}$, and S&P 500 returns $y_{SPX,t}$, according to the following empirical model.

$$y_{WTI,t} = \omega_{s_t} + \alpha_{s_t} y_{WTI,t-1} + \beta_{s_t} y_{OVX,t} + \gamma_{s_t} y_{USD,t} + \delta_{s_t} y_{SPX,t} + \varepsilon_t \quad (2)$$

This model Equation (2) implies that at a given moment, the behavior of returns does not correlate solely with its value a moment before, $y_{WTI,t-1}$, but also with contemporaneous changes in the OVX volatility index $y_{OVX,t}$, as well as returns on the US dollar index $y_{USD,t}$, and S&P 500 index $y_{SPX,t}$. As with the drift and auto-regressive parameters in Equation (1), the regression coefficients β_{s_t} , γ_{s_t} , and δ_{s_t} are assumed to depend on the state s_t prevailing at time t . As expressed in Equation (3), the regimes are assumed to

follow a first-order Markov chain in which the current state s_t depends solely on the state s_{t-1} prevailing one period before.

$$Pr(s_t = j | s_{t-1} = i, s_{t-2} = k, \dots, y_{WTI,t-1}, y_{WTI,t-2}, \dots) = Pr(s_t = j | s_{t-1} = i) = p_{ij} \quad (3)$$

The state s_t governing the return dynamics is a random variable that is not observed directly but can be inferred from the observed behavior of returns. Equation (3) assumes that the probability of regime prevalence or regime shift p_{ij} depends on past observations only through the most recent state, s_{t-1} . It is possible to examine the Markov-regime switches with n -states ($n \geq 2$) for the sake of easier exposition. The model Equation (2) is available to estimate a number of unobserved states ranging from $n = 2, \dots, 5$ for a full sample period, as well as the three subperiods A, B, and C. It is noted that assuming that there is a two-state factor in the inner dynamics of WTI returns reflecting the markets' uncertainty, which is associated with a regime of lower volatility and a regime of higher volatility, empirical modeling sets up a two-state Markov chain to examine whether there is a difference in the correlation between WTI and OVX in each regime based on an economic perspective in this paper. The matrix form of the two-state Markov chain can be expressed according to Equation (4), where all elements are non-negative, and the sums of elements in each row are equal to unity.

$$\Pi = \begin{bmatrix} p_{11} & p_{21} \\ p_{12} & p_{22} \end{bmatrix} \quad (4)$$

With reference to Equation (2), the parameters required in describing the regime probability are represented by the state-dependent drift ω_{s_t} , auto-regressive coefficient α_{s_t} and regression coefficients associated with the explanatory variables for the volatility index β_{s_t} , equity returns γ_{s_t} , and dollar index returns δ_{s_t} . The regime probability also depends on the variance σ^2 of the Gaussian distributed error terms ε_t , and on the transition probabilities p_{11} and p_{22} . Thus, the probability of switch from regime $s = i$ at time $t - 1$ to $s = j$ at time t can be expressed as $p_{ij} = 1 - p_{ii}$. A permanent shift from one regime to another would be reflected by a transition probability of unity but given the imperfect predictability of regime-switching events, it is more plausible that $p_{22} < 1$.

Assuming that past observations of the return series are available at time t in the set of information $\Omega_t = \{y_t, y_{t-1}, \dots, y_1, y_0\}$, and given the vector of regression parameters $\vartheta = (\omega_s, \alpha_s, \sigma, p_{11}, p_{22})'$, it is possible to infer, at time t , the conditional probabilities $\psi_{j,t}$ for $j = 1, 2$, according to Equation (5).

$$\psi_{j,t} = Pr(s_t = j | \Omega_t; \vartheta) \quad (5)$$

The inferred probabilities can be estimated, following Hamilton (1989, 1994), as the by-product of an iterative process, similar to a Kalman filter algorithm which predicts future states based on input from past estimators using $\psi_{i,t-1} = Pr(s_{t-1} = i | \Omega_{t-1}; \vartheta)$ for $i = 1, 2$. The iterative process is based on the density functions $\phi_{j,t}$, which can be expressed for the two-state Markov chain according to the following Equation (6).

$$\phi_{j,t} = f(y_{WTI,t} | s_t = j; \Omega_{t-1}; \vartheta) = \frac{1}{\sigma\sqrt{2\pi}} \exp \left\{ -\frac{(y_{WTI,t} - \hat{y}_{WTI,t})^2}{2\sigma^2} \right\} \quad (6)$$

where $\hat{y}_{WTI,t} = \omega_{s_t} + \alpha_{s_t}y_{WTI,t-1} + \beta_{s_t}y_{OVX,t} + \gamma_{s_t}y_{USD,t} + \delta_{s_t}y_{SPX,t}$ and the quadratic terms $(y_{WTI,t} - \hat{y}_{WTI,t})^2$ represent the squared errors ε_t^2 . Equation (7) expresses the conditional density function of return observation y_{WTI} at time t , which can be estimated from the joint density of returns and state variable:

$$f(y_{WTI,t} | \Omega_{t-1}; \vartheta) = \sum_i \sum_j p_{ij} \psi_{i,t-1} \phi_{j,t} \quad (7)$$

Following Hamilton (1989), the unknown vector of the regression model parameters $\hat{\theta}$ can be obtained using the maximum likelihood estimates of the transition probabilities according to Equation (8):

$$\hat{p}_{ij} = \frac{\sum_{t=2}^T Pr(s_t = j, s_{t-1} = i | \Omega_T; \hat{\theta})}{\sum_{t=2}^T Pr(s_{t-1} = i | \Omega_T; \hat{\theta})} \quad (8)$$

Given the starting values of the vector of parameters $\hat{\theta}_0$, the iterative process generates new sets of coefficients for the drift, autoregressive terms, and explanatory variables, as well as new estimates of the residual variance and transition probabilities. Under the assumption that the Markov chain is ergodic, it is possible to use the unconditional probabilities of Equation (5) expressed as $\psi_{i,0} = Pr(s_0 = i) = (1 - p_{jj}) / (2 - p_{ii} - p_{jj})$. The maximization of the sample conditional log likelihood, expressed in Equation (9) by numerical optimization, with iterative computations resulting in convergence toward the Maximum Likelihood (ML) estimates.

$$\log f(y_{WTI,1}, y_{WTI,2}, \dots, y_{WTI,T} | y_{WTI,0}; \theta) = \sum_{t=1}^T \log f(y_{WTI,t} | \Omega_{t-1}; \theta) \quad (9)$$

Finally, it is noted that there is a growing body of literature that addresses several issues in the ML estimation of Markov-regime switching models. For instance, Diebold et al. (1994) and Filardo (1994) examine regime-switching models where the transition probabilities are not constant as in the Hamilton study (1989) but time-varying in order to allow for the underlying fundamentals and exogenous variables to be included in the mechanism of transition between states. Additionally, Harris (1999) proposes a Bayesian Markov Chain Monte Carlo estimation of regime-switching vector autoregressions. An endogenous Markov regime-switching model was proposed by Kim et al. (2008) by relaxing the assumption that the state variable governing regime shifts is exogenous. A more recent study by Pouzo et al. (2022) examines the consistency of ML estimation with covariate-dependent transition probabilities. Thus, given the extensive interest in econometric models capable of capturing abrupt changes in economic cycles and financial time series, the estimation of standard Markov-regime switching models for the WTI futures returns may shed some light on random breaks in the inner dynamics of WTI futures returns and non-linear relationship with volatility expectations as well as equity and currency returns.

4. Empirical Evidence

4.1. Data Description and Distributional Properties

The empirical analysis is, as noted above, based on the daily time-series data for the WTI oil futures market, its related OVX volatility index, the U.S. dollar index, and the S&P 500 equity index. The sample observations obtained from the Thomson Reuters database span the time period from July 2008 to December 2021. The empirical analysis is based on the available database of the time series. The starting date of the sample period coincides with the CBOE's official release of the OVX index in 2008. It also established December 2021 as the end date of the sample period because of a different event, the progression of Ukraine by Russia, starting in 2022. It partially covers significant periods of economic uncertainty caused by the U.S. credit crisis in 2007–09, as well as the ongoing economic and healthcare crises starting in late 2019. The focus is placed in particular on the implications of the COVID-19 healthcare crisis for the inner dynamics and correlation structure changes of the WTI futures markets; therefore, the sample observations are divided into subperiod A from January 2018 to December 2019 and subperiod B from January 2020 to December 2021. The reason for using January 2020 as the point of departure is based on the WHO report that a novel coronavirus was identified in late 2019, and an emergency system was put in place to deal with a pandemic that occurred in January 2020. Additionally, the subperiod C from July 2008 to June 2010 in order to examine the impact of the U.S. government bailout, the

Federal Reserve’s Quantitative Easing announcement in response to the worsening credit crisis, and historical losses in the equity market.

It appears from the upper-left side of Figure 1 that the WTI futures prices precipitously dropped at the end of 2008 in association with the U.S. financial crisis, but the successive rebounds over the entire sample period in 2009, 2016, and 2020 have failed to regain the pre-crisis levels. Of particular interest is the historic fall on 20 April 2020 of futures prices into negative territory and significant negative returns in response to perceptions of heightened economic uncertainty stemming from the disease outbreak. The historical futures prices are also associated with a significant jump in expected volatility, as exhibited in the upper-right side of Figure 1. Judging from the typically lower scales of returns on the U.S. dollar index and S&P 500 index reported in the lower-left and right sides of Figure 1, respectively, it appears that the currency and financial markets are relatively less volatile than the energy markets. Both return series tend to exhibit more fluctuations in the earlier part of the sample period and a sudden surge in association with negative pricing of crude oil futures in 2020, but there is a clear tendency for both indices to increase in more recent years.

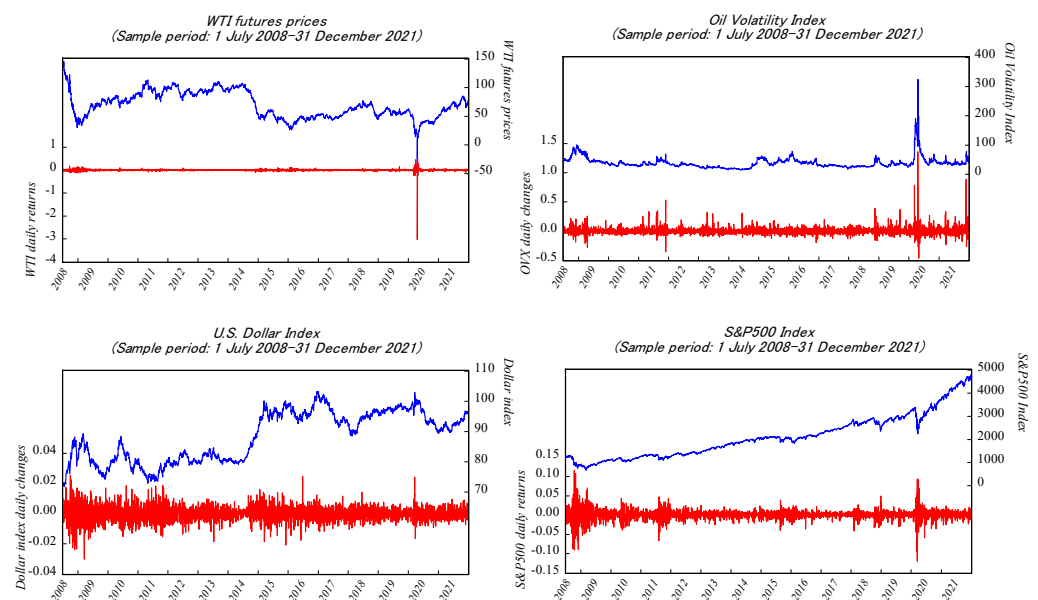


Figure 1. The behavior of price levels and returns in WTI futures, OVX, dollar index, and equity markets.

Table 1 summarizes the distributional properties of these stochastic variables. The WTI futures returns are found to be negative on average and possess more volatility than other variables. The daily changes in the OVX index tend to be positive, as volatile as the associated futures returns, and skewed to the right. In contrast, the returns on both the dollar index and S&P 500 index are found to be skewed to the left and positive on average. All time series are also found to exhibit excessive kurtosis, which is indicative of heavy-tailed distributions. Judging from the ADF test statistics, the time series are all found to be rejected stationary over the total sample period.

4.2. Model Estimation Results

The estimation results of Equation (2) with a two-state Markov chain for the full sample period are reported in Table 2. It is noted that the estimated drift term ω_1 in Markov-regime 1 is found to be statistically insignificant. It is also associated with an insignificant but negative autoregressive coefficient α_1 , which suggests a tendency for mean reversion. The crude oil futures returns are also governed by a strong negative correlation with changes in the OVX volatility index. With significance at the one-percent level, the negative sign of the regression coefficient β_1 implies that an increase in expected volatility is associated with

diminishing oil futures returns. Given the significantly negative regression coefficients γ_1 and δ_1 , the dynamics of futures returns are also found to be sensitive to changes in the U.S. dollar index and equity valuation.

Table 1. Distributional Properties.

Distributional Moments	Mean	Std. Dev.	Skewness	Kurtosis	Jarque Bera	ADF Test
WTI returns	−0.0007	0.062728	−33.5240	1569.688	3.61×10^8	−43.544 ***b
OVX daily changes	0.00174	0.063033	4.9582	87.09921	1,052,944.0	−61.760 ***a
Dollar index daily changes	0.00009	0.004761	−0.0492	5.870737	1211.493	−58.464 ***b
S&P 500 returns	0.00045	0.012766	−0.2904	17.14844	29,442.39	−68.718 ***a

Notes: The sample period of daily observation runs from 1 July 2008, to 31 December 2021. Significance at the 1% level is denoted by asterisks *** under MacKinnon (1996)’s one-sided probability values. The stationarity of time series is estimated with the Augmented Dickey–Fuller methodology using the intercept only and with neither intercept nor trend terms as denoted by superscripts a and b respectively. Jarque–Bera statistics for normal tests are distributed as χ^2 on the null.

Table 2. Markov-regime switching model results (Full estimated period).

Model Parameters	Full Period (July 2008–December 2021)	
	Markov-Regime 1	Markov-Regime 2
ω	−0.0001 (0.6697)	0.0174 (0.2269)
α	−0.0196 (0.1708)	0.1592 *** (0.0045)
β	−0.1452 *** (0.0000)	−0.9643 *** (0.0000)
γ	−0.8659 *** (0.0000)	−1.4004 (0.4340)
δ	0.3820 *** (0.0000)	−0.6599 (0.2359)
$\text{Log}(\sigma)$	−4.1048 *** (0.0000)	−1.6497 *** (0.0000)
Log likelihood	9124.611	
AIC	−4.9880	
Hypothesis Tests		
$\omega_1 = \omega_2$	1.4408 (0.2300)	
$\alpha_1 = \alpha_2$	9.5865 *** (0.0020)	
$\beta_1 = \beta_2$	76.7273 *** (0.0000)	
$\gamma_1 = \gamma_2$	0.0259 (0.8723)	
$\delta_1 = \delta_2$	3.2807 * (0.0701)	

Notes: The estimated Markov-regime Switching model is represented by equation: $y_{WTI,t} = \omega_{s_t} + \alpha_{s_t}y_{WTI,t-1} + \beta_{s_t}y_{OVX,t} + \gamma_{s_t}y_{USD,t} + \delta_{s_t}y_{SPX,t} + \varepsilon_t$. The sample period of daily observation runs from 1 July 2008 to 31 December 2021. Significance at 1 and 10% level is denoted by *** and *, respectively. The hypothesis tests for equal coefficients are based on the Wald test following the χ^2 distribution. Figures in round brackets represent probability values.

In contrast, returns governed by the Markov-regime 2 are characterized by positive but insignificant drift ω_2 , positive autoregressive coefficient α_2 that implies long memory rather than mean reversion, and a negative correlation with changes in volatility expectations.

However, it seems that the return dynamics are not sensitive to contemporaneous variations in the U.S. dollar index or stock prices. It is also noted that both regimes are associated with significant volatility estimates, though Regime 2 seems to exhibit a relatively higher level of fluctuations. The χ^2 -distributed Wald tests of hypothesis for equal regression parameters indicate that it is possible to distinguish regimes based on different autoregressive terms, as well as the structure of return correlation with changes in volatility index and equity returns. Indeed, despite the opposite signs, the null hypothesis of equal drifts cannot be rejected as both drifts are found to be insignificantly different across regimes despite the opposite signs. Conversely, the oil futures returns are characterized by mean reversion in Regime 1; they tend to exhibit long memory in Regime 2. The Wald test results indicate that it is difficult to distinguish between regimes on the basis of the relationship between returns on oil futures and the U.S. dollar index. However, the difference between the regression coefficients associated with equity returns is found to be significant only at the ten-percent level. It is the extent to which the correlation with changes in the OVX index seems to be most prominent in distinguishing between regimes. Thus, a transition from a regime of lower volatility to one of higher volatility, i.e., from Regime 1 to Regime 2, is accompanied by an increase in the significance of negative correlation with volatility expectations. This new evidence sheds light on the shifting expectations of market participants about future levels of uncertainty and the need to consider the dynamics of investors' fear as a significant determinant of the return-generating process.

It is possible to examine the frequency of switches between the latent states based on the probability of Regime 1, which is shown in Figure 2, together with the time series of daily WTI futures prices. It is clear that it is Regime 1 of lower volatility that tends to dominate over long durations, but there are frequent shifts to Regime 2 at the beginning of the total sample period from August 2008 to April 2009 and in association with periods of persistent price falls from January 2016 to May 2016 as well as with the precipitous decrease in futures prices in April 2020 from March 2020 to June 2020. This evidence is partly consistent with the Chow test of a structural break in the same model Equation (2), which indicates the existence of a single break dated 25 December 2019, at the 5% significance level (F-Statistic 173.992, Bai–Perron critical value 18.23). This result seems to be consistent with the empirical evidence from the estimated Markov-regime switching model, which suggests frequent regime shifts in relation to the onset of the disease outbreak in December 2019, as well as the subsequent government responses and market reactions over the crisis period from March 2020 to June 2020. Thus, it seems that the dynamics of oil futures returns are responsive to the policy responses of the U.S. government to the onset of the credit crisis, as well as to the heightened levels of uncertainty about the global economy in association with the disease outbreak. Market perceptions of higher economic uncertainty in association with major events are conducive to abrupt shifts from a regime characterized by long memory rather than mean reversion and stronger rather than the weaker negative correlation with the forward-looking measure of oil volatility. Indeed, the higher the perceived levels of uncertainty in the crude oil markets, the stronger the negative correlation between WTI futures returns and changes in volatility expectations.

Thus, the graphical evidence from Figure 2 suggests that regime shifts are more likely to occur frequently during periods of decreasing WTI futures prices and higher economic uncertainty. In order to examine the non-linear dynamics prior to and during the disease outbreak, the model Equation (2) with a two-state Markov chain is estimated for both subperiod A from January 2018 to December 2019 and subperiod B from January 2020 to December 2021. Judging by the results reported in Table 3 for subperiod A, it appears that futures returns are governed by two distinct regimes characterized by different levels of volatility. Regime 1 is associated with a statistically insignificant drift and autoregressive term, as well as an insignificant relationship with equity returns, but with negative correlations with changes in the OVX volatility index and in the dollar index. Though the latter is significant only at the ten-percent level, it is found to be insignificant under the alternative regime. Indeed, futures returns are characterized, under Regime 2, by positive

drift and positive correlation with equity returns, albeit significant only at the ten-percent and five-percent levels, respectively.

Table 3. Results of Subperiod A and B estimated with Markov-regime switching modeling.

Model Parameters	Subperiod A (January 2018–December 2019)		Subperiod B (January 2020–December 2021)	
	Regime 1	Regime 2	Regime 1	Regime 2
ω	−0.0009 (0.2648)	0.0056 * (0.0811)	0.0014 * (0.0835)	0.0135 (0.7792)
α	0.0145 (0.7738)	−2.4246 ** (0.0153)	0.0014 (0.9667)	0.1383 (0.1553)
β	−0.2205 *** (0.0000)	0.2100 *** (0.0000)	−0.1802 *** (0.0000)	−1.2691 *** (0.0000)
γ	−0.4526 * (0.0671)	−0.1772 (0.8440)	−0.5188 ** (0.0379)	−2.3555 (0.7371)
δ	0.0913 (0.3084)	0.8216 ** (0.0488)	0.2888 *** (0.0003)	−1.9662 (0.1892)
$\text{Log}(\sigma)$	−4.3179 *** (0.0000)	−3.7329 *** (0.0000)	−4.1164 ** (0.0000)	−1.1437 *** (0.0000)
Log Likelihood	1403.674		1219.407	
AIC	−5.324421		−4.609587	
Hypothesis tests				
$\omega_1 = \omega_2$	3.4875 * (0.0618)		0.0636 (0.8009)	
$\alpha_1 = \alpha_2$	4.3483 ** (0.0370)		1.7564 (0.1851)	
$\beta_1 = \beta_2$	84.6471 *** (0.0000)		38.8884 *** (0.0000)	
$\gamma_1 = \gamma_2$	0.0783 (0.7796)		0.0684 (0.7936)	
$\delta_1 = \delta_2$	2.6231 (0.1053)		2.2305 (0.1353)	

Notes: The estimated Markov-regime Switching model is represented by the equation: $y_{WTL,t} = \omega_{s_t} + \alpha_{s_t} y_{WTL,t-1} + \beta_{s_t} y_{OVX,t} + \gamma_{s_t} y_{USD,t} + \delta_{s_t} y_{SPX,t} + \varepsilon_t$. The sample period of daily observation runs from 1 January 2018 to 31 December 2021. Significance at the 1, 5, and 10% levels is denoted by ***, **, and *, respectively. The hypothesis tests for equal coefficients are based on the Wald test following the χ^2 distribution. Figures in round brackets represent probability values.

There is also evidence that Regime 2 implies mean reversion and a positive correlation with changes in the OVX index. In contrast to Regime 1 and to both regimes for the full sample period, the significance of the β_2 coefficient suggests that futures returns tend instead to rise in association with increasing uncertainty. This is consistent with the evidence from Aboura and Chevaller (2013), who found a positive correlation between changes in OVX levels and oil prices in association with the onset of the U.S. credit crisis. Judging from the tests of null hypotheses for equal regression coefficients, it appears that the Markov regimes can be distinguished not so much on the basis of differences in the correlations with currency and equity returns as differences between autoregressive terms and correlation with volatility expectations.

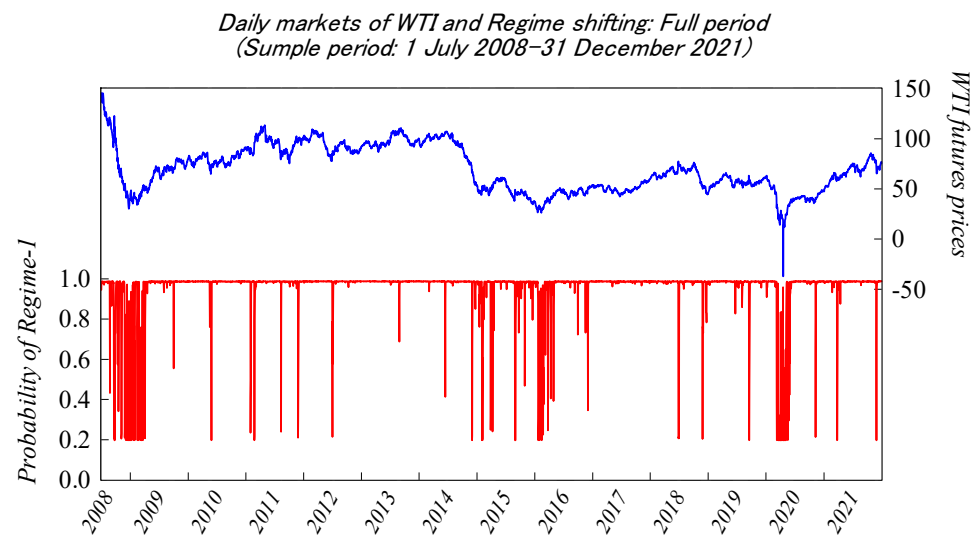


Figure 2. Probability of Regime 1 (Full sample period July 2008–December 2021).

With respect to the estimation results for subperiod B, also reported in Table 3, there is clear evidence that the behavior of oil futures returns is likely to be governed by two Markov regimes that tend to differ only on the basis of weaker and stronger correlation with changes in the OVX index. Indeed, the drift terms for both regimes are associated with positive signs but statistically significant only for Regime 1 at the ten-percent level. Additionally, the autoregressive terms are found to be statistically insignificant for both regimes. The relationship of futures returns with the dollar index returns is found to be negative at the five-percent level for Regime 1 but insignificant for Regime 2. Similarly, the regression coefficient δ_s reflecting the sensitivity of futures returns to changes in equity valuation is found to be positive under Regime 1 but insignificant under the alternative regime. It is clear that only the regression coefficients β_s are found to be negative and significant at the one-percent level under both regimes. Thus, a shift from Regime 1 is likely to be accompanied by a strong increase in sensitivity to changes in volatility expectations under Regime 2.

Judging by the estimated probability values reported in Figure 3, it is Regime 1 that seems to predominate over subperiod A. The regime switches are not likely to take place as the decrease in the likelihood of Regime 1 remains above the fifty-percent threshold probability value. Thus, it is the Regime with a strong negative correlation with volatility expectations that is more likely to prevail. Similarly, the evidence from Figure 4, which reports the Regime 1 probability for subperiod B, suggests that there are frequent regime shifts in association with the precipitous fall in futures prices below zero. A shift toward Regime 2 implies that future returns are governed by a stronger negative correlation with changes in the OVX index. This suggests that higher levels of volatility expectations are conducive to even lower futures returns. Apart from the short period of negative future pricing, it is Regime 1 that tends to prevail with a weaker but still significantly negative correlation with volatility expectations.

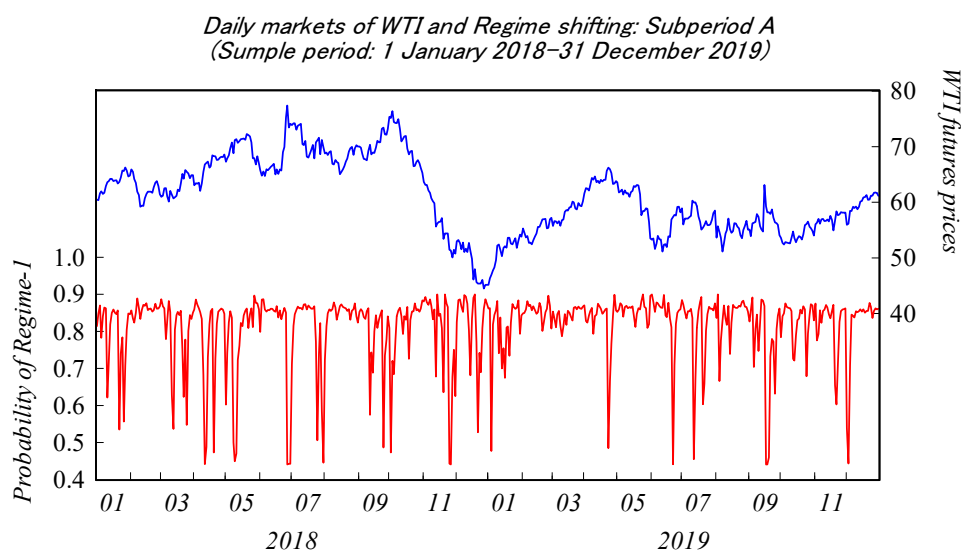


Figure 3. Probability of Regime 1 (Subperiod A—January 2018–December 2021).

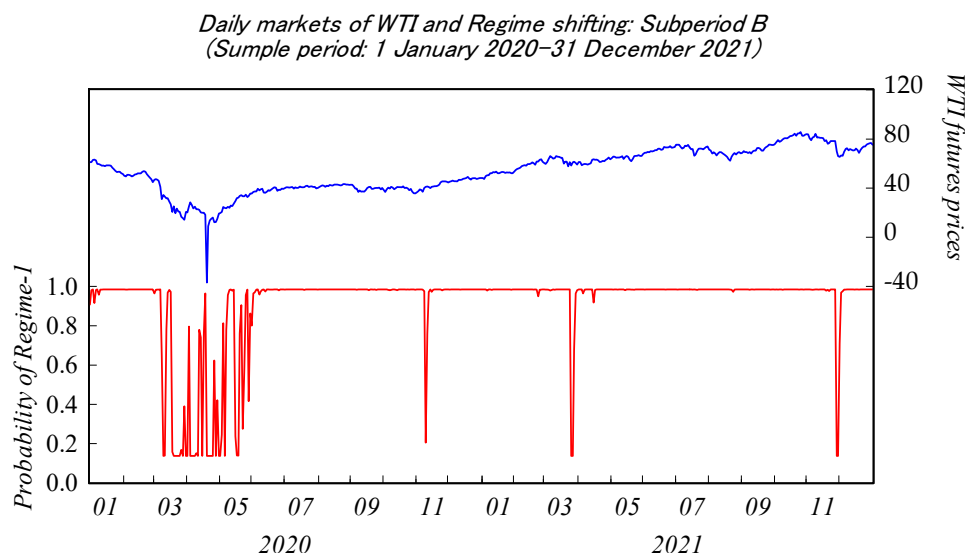


Figure 4. Probability of Regime 1 (Subperiod B—January 2020–December 2021).

In order to further examine the robustness of two-state Markov-regime switching models to changes in the sample period, the focus is made on subperiod C from July 2008 to June 2010, which may be in part reflective of the effects of the U.S. credit crisis on the inner dynamics of oil futures returns and their relationship with volatility expectations and alternative asset markets. The evidence from Table 4, which reports the estimation of the regime-switching model with a two-state Markov chain, indicates that both regimes are characterized by statistically insignificant drifts ω_s and autoregressive terms α_s . However, futures returns are likely to be governed by strong negative correlation with dollar index returns γ_s and positive correlations with equity returns δ_s under both regimes. In contrast, a shift from Regime 1 to Regime 2 is likely to result in a strong negative correlation between futures returns and volatility expectations fading away. However, the aggregate evidence from Wald tests of the null hypothesis of equal coefficients suggests that it is difficult to distinguish between these Markov regimes and that, given their similar properties, it is more likely that the two regimes collapse into a single one with significantly negative sensitivity to dollar index returns $\gamma < 0$, positive sensitivity to $\delta > 0$, and a more likely negative correlation with changes in the OVX index $\beta \leq 0$. Thus, it is clear from the estimation results for subperiod C that periods of financial instability can weaken the correlation of futures

returns with volatility expectations. Under higher levels of uncertainty, an increase in oil volatility expectations may not be conducive to diminishing futures returns. It is rather the linkage between oil futures returns and financial markets that gains more significance. Increasing oil futures returns are more likely to result from lower dollar valuation and higher equity.

Table 4. Markov-regime switching model results (Subperiod C estimated period).

Model Parameters	Subperiod C (July 2008–June 2010)	
	Markov-Regime 1	Markov-Regime 2
ω	0.0006 (0.4797)	−0.0008 (0.8519)
α	0.0621 (0.1126)	−0.0018 (0.8062)
β	−0.1402 *** (0.0000)	−0.0958 (0.1062)
γ	−1.7029 *** (0.0000)	−1.1439 ** (0.0169)
δ	0.3553 *** (0.0000)	0.4894 *** (0.0006)
$\text{Log}(\sigma)$	−4.1663 *** (0.0000)	−2.9762 *** (0.0000)
Log likelihood	1220.131	
AIC	−4.6212	
Hypothesis Tests		
$\omega_1 = \omega_2$	0.1054 (0.7454)	
$\alpha_1 = \alpha_2$	0.8807 (0.3480)	
$\beta_1 = \beta_2$	0.3969 (0.5287)	
$\gamma_1 = \gamma_2$	1.0965 (0.2950)	
$\delta_1 = \delta_2$	0.5650 (0.4523)	

Notes: The estimated Markov-regime Switching model is represented by the equation: $y_{WTI,t} = \omega_{s_t} + \alpha_{s_t} y_{WTI,t-1} + \beta_{s_t} y_{OVX,t} + \gamma_{s_t} y_{USD,t} + \delta_{s_t} y_{SPX,t} + \varepsilon_t$. The sample period of daily observation runs from 1 July 2008 to 30 June 2010. Significance at the 1 and 5% levels is denoted by *** and **, respectively. The hypothesis tests for equal coefficients are based on the Wald test following the χ^2 distribution. Figures in round brackets represent probability values.

4.3. Robustness Checks

This paper assumes that there is two-state, lower volatility under bullish markets and higher volatility under bearish markets, in the WTI futures market that reflects market uncertainty and conducts an empirical analysis using the Markov-regime switching model (a non-linear model) to determine whether the correlation between WTI returns and OVX daily change performs with a two-state view of the market depending on the fluctuation of markets. Furthermore, it is possible to consider empirical analysis from a different perspective without taking into account the inner market states. The approach uses a linear model to analyze the correlation between WTI returns and OVX daily fluctuations before and after the structural change date, respectively. Thus, in this section, it is used the autoregressive distributed lag (ARDL) model, one of the linear models, to test whether the empirical analysis demonstrated in this paper is robust. This section also examines

whether especially WTI and OVX have a cointegration relationship in each period before and after the structural change using the bounds-testing approach of Pesaran et al. (2001).

In the first place, an explanation of the empirical model used in the linear model estimation is provided. Equation (10) is based on the general ARDL model and incorporates variables similar to those in Equation (2) used in this paper.

$$y_{WTI,t} = \omega + \sum_{k=1}^m \alpha_k y_{WTI,t-k} + \sum_{k=0}^m \beta_k y_{OVX,t-k} + \sum_{k=0}^m \gamma_k y_{USD,t-k} + \sum_{k=0}^m \delta_k y_{SPX,t-k} + \varepsilon_t \quad (10)$$

where the disturbance terms are white noise distributed with $\varepsilon_t \sim i.i.d.N(0, \sigma^2)$. m is lag up to 1, 2, . . . , 8, where m is selected as the optimal lag length by the AIC. Then, in order to estimate whether all variables, WTI returns, OVX daily changes, dollar index daily changes, and S&P 500 index returns, have a long-run relationship, an extension is made from Equation (10) to Equation (11), following the method derived by Pesaran et al. (2001). Equation (11) can be expressed as follows:

$$\Delta y_{WTI,t} = a_0 + \sum_{i=1}^m a_{1,i} \Delta y_{WTI,t-i} + \sum_{i=0}^m a_{2,i} \Delta y_{OVX,t-i} + \sum_{i=0}^m a_{3,i} \Delta y_{USD,t-i} + \sum_{i=0}^m a_{4,i} \Delta y_{SPX,t-i} + a_5 y_{WTI,t-1} + a_6 y_{OVX,t-1} + a_7 y_{USD,t-1} + a_8 y_{SPX,t-1} + \varepsilon_t \quad (11)$$

where Δ is the first difference operator, $a_0, a_1, a_2, \dots, a_7$, and a_8 are parameters, m is the optimal lag length to be used for estimation selected by AIC. The bounds-testing approach is based on the F-Statistic and is the first of the ARDL cointegration methods. The null hypothesis test of no cointegration, ($H_0 : a_5 = a_6 = a_7 = a_8 = 0$), will be performed by Equation (11). Following Pesaran et al. (2001), it is computed two sets of critical values for a given significance level. One set assumes that all variables are $I(0)$, and the other set assumes all variables are $I(1)$. There are three cases that will be obtained. In case one, if the estimated F-Statistic exceeds the upper critical value, the null hypothesis is rejected. In case two, if the estimated F-Statistic is between the upper critical value and lower critical value, then the testing becomes inconclusive. In case three, if the estimated F-Statistic is below the lower critical value, it suggests no cointegration among all variables.

As noted in Section 4.2, the result from the Chow test for structural breaks indicates that the date on which the existence of structural break is 25 December 2019, and it can be regarded as approximately equal to the base point of dividing subperiods A and B. Therefore, it is performed estimation using Equations (10) and (11) in sub-periods A and B to verify the change in the correlation between WTI returns and OVX daily changes before and after a structural change and to test whether there is a long-run relationship between all variables.

Table 5 summarizes the results estimated in subperiod A using Equation (10), which reflects the optimal lag length selected by AIC, and the results of the bounds test for examining the long-run relationship between all variables using Equation (11). The estimated result with Equation (10) for subperiod A indicates that contemporaneous to three periods earlier, OVX daily changes are weakly negatively correlated with WTI returns. The absolute value of the t-value is the largest for the contemporaneous OVX daily change, suggesting that the contemporaneous period of OVX daily changes is more influential than the other lags to WTI returns where time t . Additionally, the results of the bounds testing with Equation (11) indicate that the estimated F-Statistic exceeds the upper critical value, which means that it has rejected the no levels of relationship at the 1% significance level (F-Statistic), suggesting there is a cointegration relationship between all variables in subperiod A.

Table 5. Results of Subperiod A estimated with ARDL modeling and the bounds test.

Model Parameters [ARDL (5,3,2,0)]	Subperiod A (January 2018–December 2019)	
	Coefficient	t-Statistic
ω	0.0007	0.8380
α_{t-1}	−0.1567 ***	−3.5950
α_{t-2}	−0.0658	−1.4974
α_{t-3}	−0.0814 *	−1.8818
α_{t-4}	0.0582	1.4148
α_{t-5}	0.0733 *	1.7949
β_t	−0.1010 ***	−6.2670
β_{t-1}	−0.0648 ***	−4.0226
β_{t-2}	−0.0360 **	−2.2070
β_{t-3}	−0.0317 *	−1.9481
γ_t	−0.2385	−0.9302
γ_{t-1}	−0.5504 **	−2.1447
γ_{t-2}	−0.4391 *	−1.7066
δ_t	0.2806 ***	2.9526
Log Likelihood	1340.482	
AIC	−5.0823	
F-Bounds Test (At the 1% significance level)		
F-Statistic	I(0)	I(1)
29.9258	3.65	4.66

Notes: The estimated ARDL model is represented by Equation (10): $y_{WTI,t} = \omega + \sum_{k=1}^m \alpha_k y_{WTI,t-k} + \sum_{k=0}^m \beta_k y_{OVX,t-k} + \sum_{k=0}^m \gamma_k y_{USD,t-k} + \sum_{k=0}^m \delta_k y_{SPX,t-k} + \varepsilon_t$. The sample period of daily observation runs from 1 January 2018 to 31 December 2019. Significance at the 1, 5, and 10% levels is denoted by ***, **, and *, respectively. The hypothesis test of cointegration for all variables is based on the bounds testing with Equation (11): $\Delta y_{WTI,t} = a_0 + \sum_{i=1}^m a_{1,i} \Delta y_{WTI,t-i} + \sum_{i=0}^m a_{2,i} \Delta y_{OVX,t-i} + \sum_{i=0}^m a_{3,i} \Delta y_{USD,t-i} + \sum_{i=0}^m a_{4,i} \Delta y_{SPX,t-i} + a_5 y_{WTI,t-1} + a_6 y_{OVX,t-1} + a_7 y_{USD,t-1} + a_8 y_{SPX,t-1} + \varepsilon_t$.

On the other hand, Table 6, which reports the results estimated in subperiod B using Equation (10) reflecting the optimal lag length selected by AIC, shows that only the contemporaneous OVX daily change is strongly negatively correlated with the WTI return, and the absolute value of the t value is larger than the other variables, the dollar index daily change and S&P 500 returns, suggesting that the contemporaneous period of OVX daily changes is more influential than the other lags and variables to WTI returns where time t. Table 6 also reports the results of the bound test with Equation (11). As in subperiod A, the result indicates that the estimated F-Statistic exceeds the upper critical value, which means that it has rejected the no levels of relationship at the 1% significance level (F-Statistic), suggesting there is also a cointegration relationship between all variables in subperiod B.

From the estimation results for subperiods A and B using the linear ARDL model, it can be observed that the correlation between WTI returns and OVX daily changes changed with the occurrence of COVID-19, which is consistent with the empirical results in this paper. In addition, the fact that a cointegration relationship is established in both subperiods from the results of using Pesaran et al.’s (2001) bounds test suggests that even with structural changes, there is the long-run relationship, which is an important point in the crude oil market. Thus, the empirical results of this paper can be regarded as robust because the change in the correlation between WTI returns and OVX daily changes before and after the structural change can also be observed in the estimation using the linear model and is consistent with the results of this paper.

Table 6. Results of Subperiod B estimated with ARDL modeling and the bounds test.

Model Parameters [ARDL (4,0,4,1)]	Subperiod B (January 2020–December 2021)	
	Coefficient	t-Statistic
ω	0.0037	0.7861
α_{t-1}	0.2494 ***	7.2207
α_{t-2}	−0.1279 ***	−3.6187
α_{t-3}	0.0400	1.1396
α_{t-4}	0.0577 *	1.7318
β_t	−0.8823 ***	−18.6752
γ_t	−1.2195	−0.9081
γ_{t-1}	0.5592	0.4318
γ_{t-2}	−1.5832	−1.2145
γ_{t-3}	−3.7311 ***	−2.8466
γ_{t-4}	3.2288 **	2.4446
δ_t	−1.4937 ***	−4.2291
δ_{t-1}	−1.5759 ***	−4.5938
Log Likelihood	426.7656	
AIC	−1.5823	
F-Bounds Test (At the 1% significance level)		
F-Statistic	I(0)	I(1)
109.6350	3.65	4.66

Notes: The estimated ARDL model is represented by Equation (10): $y_{WTI,t} = \omega + \sum_{k=1}^m \alpha_k y_{WTI,t-k} + \sum_{k=0}^m \beta_k y_{OVX,t-k} + \sum_{k=0}^m \gamma_k y_{USD,t-k} + \sum_{k=0}^m \delta_k y_{SPX,t-k} + \varepsilon_t$. The sample period of daily observation runs from 1 January 2020 to 31 December 2020. Significance at the 1, 5, and 10% levels is denoted by ***, **, and *, respectively. The hypothesis test of cointegration for all variables is based on the bounds testing with Equation (11): $\Delta y_{WTI,t} = a_0 + \sum_{i=1}^m a_{1,i} \Delta y_{WTI,t-i} + \sum_{i=0}^m a_{2,i} \Delta y_{OVX,t-i} + \sum_{i=0}^m a_{3,i} \Delta y_{USD,t-i} + \sum_{i=0}^m a_{4,i} \Delta y_{SPX,t-i} + a_5 y_{WTI,t-1} + a_6 y_{OVX,t-1} + a_7 y_{USD,t-1} + a_8 y_{SPX,t-1} + \varepsilon_t$.

5. Conclusions

The present study provides new empirical evidence on the stochastic behavior of energy futures returns based on the estimation of Markov-regime switching models. Given the nature of energy markets, the demand and supply functions are intrinsically linked with the real economy and perceptions of economic uncertainty, but there is growing literature about stronger linkages with the financial economy as well. The focus of this paper is placed on the empirical issue of whether the inner dynamics and correlation structures of oil futures with alternative asset markets are governed by different regimes that reflect changes in the underlying demographic, macroeconomic and social conditions. The empirical evidence suggests that oil futures returns tend to be governed by different Markov regimes, which invariably exhibit a negative correlation with volatility expectations which reflect the shifting fear factor dynamics. Thus, market perceptions of heightened economic uncertainty reflected by increased volatility expectations are conducive to diminishing futures returns, irrespective of the prevailing regime.

The non-linear dynamics of oil futures returns can be altered, however, by significant events such as the onset of financial crises and disease outbreaks, as well as government policy responses. There is, indeed, evidence that the economic lockdowns in response to the disease outbreak have the potential to increase the likelihood of Markov regimes with a more pronounced negative correlation of futures returns with changes in expected volatility. Additionally, periods of financial instability, such as the U.S. credit crisis, may sever the relationship of oil futures returns with volatility expectations and strengthen their linkages with currency fluctuations and equity valuation. Thus, Markov-regime switching models have the capacity to capture changes in the underlying fundamentals and provide some insights into the changing inner dynamics, such as the propensity for mean reversion and a long memory for economic shocks that tend to decay over a longer time. Regrettably, this paper’s shortcomings include the fact that the Markov-regime switching model does not resolve the issues of strict simultaneity and endogeneity between WTI and OVX, which is an

issue for future research. Further research may shed light on the non-linear dynamics with time-varying transition probabilities, simultaneity, and jump-diffusion processes which may better account for the stochastic properties of energy prices.

Funding: This research received no external funding.

Data Availability Statement: The data presented in this paper are available upon request from the author. All time-series data used in the analysis in this paper can be obtained from the Thomson Reuters database.

Acknowledgments: I would like to express my appreciations to the three anonymous reviewers who reviewed my paper, to Kentaro Iwatsubo from Kobe University for his comments this paper and Nabil Maghrebi from Wakayama University for proofreading the English text.

Conflicts of Interest: The author declares no conflict of interest.

Notes

- ¹ It is noted that the study by Fleming et al. (1995) was performed with a volatility index based on the S&P 100 stock market index, formerly known as VIX index.
- ² The focus is also placed, as in Mencia and Sentana (2013), on the valuation of VIX derivatives, where the volatility index serves as the underlying asset for derivatives contracts.
- ³ See for instance, Kim et al. (2019), Wang and Xie (2012), Choi and Hammoudeh (2010), Mensi et al. (2013), Raza et al. (2016) and Creti et al. (2013), inter alia.

References

- Aboura, Sofiane, and Julien Chevallier. 2013. Leverage vs. feedback: Which Effect drives the oil market? *Finance Research Letters* 10: 131–41. [CrossRef]
- Baiardi, Lorenzo Cerboni, Massimo Costabile, Domenico De Giovanni, Fabio Lamantia, Arturo Leccadito, Ivar Massabó, Massimiliano Menzietti, Marco Pirra, Emilio Russo, and Alessandro Staino. 2020. The Dynamics of the S&P 500 under a crisis context: Insights from a Three-Regime Switching Model. *Risks* 8: 71.
- Baumeister, Christian, and Lutz Kilian. 2012. Real-Time Forecasts of the Real Price of Oil. *Journal of Business & Economic Statistics* 30: 326–36.
- Baumeister, Christian, and Lutz Kilian. 2015. Forecasting the Real Price of Oil in a Changing World: A Forecast Combination Approach. *Journal of Business & Economic Statistics* 33: 338–51.
- Beckmann, Joscha, Robert L. Czudaj, and Vipin Arora. 2020. The relationship between oil prices and exchange rates: Revisiting theory and evidence. *Energy Economics* 88: 104772. [CrossRef]
- Bodie, Zvi, and Victor I. Rosansky. 1980. Risk and Return in Commodity Futures. *Financial Analysts Journal* 36: 27–39. [CrossRef]
- Chen, Yanhui, Kaijian He, and Lean Yu. 2015. The Information Content of OVX for Crude Oil Returns Analysis and Risk Measurement: Evidence from the Kalman Filter Model. *Annals of Data Science* 2: 471–87. [CrossRef]
- Cheng, Ing-Haw, and Wei Xiong. 2014. Financialization of Commodity Market. *Annual Review of Financial Economics* 6: 419–41. [CrossRef]
- Cheung, C. Sherman, and Peter Miu. 2010. Diversification benefits of commodity futures. *Journal of International Financial Markets, Institutions and Money* 20: 451–74. [CrossRef]
- Choi, K., and S. Hammoudeh. 2010. Volatility behavior of oil, industrial commodity and stock markets in a regime-switching environment. *Energy Policy* 38: 4388–99. [CrossRef]
- Connolly, Robert A., Chris Stivers, and Licheng Sun. 2005. Stock Market Uncertainty and the Stock-Bond Return Relation. *Journal of Financial and Quantitative Analysis* 40: 161–94. [CrossRef]
- Creti, Anna, Marc Joëts, and Valérie Mignon. 2013. On the links between stock and commodity markets' volatility. *Energy Economics* 37: 16–28. [CrossRef]
- Diebold, Francis X., Joon-Haeng Lee, and Gretchen C. Weinbach. 1994. Regime switching with time-varying transition probabilities. In *Advanced Texts in Econometrics*. Edited by Clive William John Granger and Grayham Mizon. Oxford: Oxford University Press, pp. 283–302.
- Dupoyet, Brice V., and Corey A. Shank. 2018. Oil prices implied volatility or direction: Which matters more to financial markets? *Financial Markets and Portfolio Management* 32: 275–95. [CrossRef]
- Dutta, Anupam. 2017. Oil price uncertainty and clean energy stock returns: New evidence from crude oil volatility index. *Journal of Cleaner Production* 164: 1157–66. [CrossRef]
- Federal Reserve Board. 2022. *Financial Stability Report*; Washington, DC: The United States of America Board of Governors of The Federal Reserve System, May 9.



- Ferraro, Domenico, Kenneth Rogoff, and Barbara Rossi. 2015. Can oil prices forecast exchange rates? An empirical analysis of the relationship between commodity prices and exchange rates. *Journal of International Money and Finance* 54: 116–41. [CrossRef]
- Filardo, Andrew J. 1994. Business cycle phases and their transitions. *Journal of Business and Economic Statistics* 12: 299–308.
- Fleming, Jeff, Barbara Ostdiek, and Robert E. Whaley. 1995. Predicting stock market volatility: A new measure. *The Journal of Futures Markets* 15: 265–302. [CrossRef]
- Fukasawa, Masaaki, Isao Ishida, Nabil Maghrebi, Kosuke Oya, Masato Ubukata, and Kazutoshi Yamazaki. 2011. Model-free implied volatility: From surface to index. *International Journal of Theoretical and Applied Finance* 14: 433–63. [CrossRef]
- Giot, Pierre. 2003. *The Asian Financial Crisis: The Start of a Regime Switch in Volatility*. CORE Discussion Paper, No.2003/78. Rochester: SSRN.
- Goldstein, Itay, and Liyan Yang. 2022. Commodity Financialization and Information Transmission. *The Journal of Finance* 77: 2613–67. [CrossRef]
- Gong, Xu, and Jun Xu. 2022. Geopolitical risk and dynamic connectedness between commodity markets. *Energy Economics* 110: 106028. [CrossRef]
- Gorton, Gary, and K. Geert Rouwenhorst. 2006. Facts and Fantasies about Commodity Futures. *Financial Analysts Journal* 62: 47–68. [CrossRef]
- Grima, Simon, Letife Özdemir, Ercan Özen, and Inna Romānova. 2021. The Interactions between COVID-19 Cases in the USA, the VIX Index and Major Stock Markets. *International Journal of Financial Studies* 9: 26. [CrossRef]
- Hamilton, James D. 1989. A New approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica* 57: 357–84. [CrossRef]
- Hamilton, James D. 1994. *Time Series Analysis*. Princeton: Princeton University Press.
- Hamilton, James D. 2008. *Understanding Crude Oil Prices*. National Bureau of Economic Research Working Paper, No. 14492. Cambridge: National Bureau of Economic Research.
- Harris, Glen R. 1999. Markov chain Monte Carlo estimation of regime switching vector autoregressions. *Austin Bulletin* 29: 47–79. [CrossRef]
- Jensen, Gerald R., Robert R. Johnson, and Jeffrey M. Mercer. 2000. Efficient use of commodity futures in diversified portfolios. *The Journal of Futures Markets* 20: 489–506. [CrossRef]
- Just, Małgorzata, and Krzysztof Echaust. 2020. Stock market returns, volatility, correlation and liquidity during the COVID-19 crisis: Evidence from the Markov Switching approach. *Finance Research Letters* 37: 101775. [CrossRef]
- Kim, Chang-Jin, Jeremy M. Piger, and Richard Startz. 2008. Estimation of Markov regime-switching regression models with endogenous switching. *Journal of Econometrics* 143: 263–73. [CrossRef]
- Kim, Suyi, So-Yeun Kim, and Kyungmee Choi. 2019. Analyzing oil price shocks and exchange rates movements in Korea using Markov Regime-Switching models. *Energies* 12: 4581. [CrossRef]
- Liu, Bing-Yue, Qiang Ji, and Ying Fan. 2017. Dynamic return-volatility dependence and risk measure of CoVaR in the oil market: A time-varying mixed copula model. *Energy Economics* 68: 53–65. [CrossRef]
- Lombardi, Marco, and Francesco Ravazzolo. 2013. *On the Correlation between Commodity and Equity Returns: Implications for Portfolio Allocation*. BIS Working Papers No, 420. Basel: Bank for International Settlements.
- MacKinnon, James G. 1996. Numerical Distribution Function for Unit root and Cointegration tests. *Journal of Applied Econometrics* 11: 601–18. [CrossRef]
- Maghrebi, Nabil, Moo-Sung Kim, and Kazuhiko Nishina. 2007. The KOSPI 200 Implied Volatility Index: Evidence of Regime Switches in Volatility Expectations. *Asia-Pacific Journal of Financial Studies* 36: 163–87.
- Maghrebi, Nabil, Mark J. Holmes, and Kosuke Oya. 2014. Financial instability and the short-term dynamics of volatility expectations. *Applied Financial Economics* 24: 377–95. [CrossRef]
- Mencia, Javier, and Enrique Sentana. 2013. Valuation of VIX derivatives. *Journal of Financial Economics* 108: 367–91. [CrossRef]
- Mensi, Walid, Makram Beljid, Adel Boubaker, and Shunsuke Managi. 2013. Correlations and volatility spillovers across commodity and stock markets: Linking energies, food, and gold. *Economic Modelling* 32: 15–22. [CrossRef]
- Nishina, Kazuhiko, Nabil Maghrebi, and Moo-Sung Kim. 2006. *Stock Market Volatility and the Forecasting Accuracy of Implied Volatility Indices*. Discussion Papers in Economics and Business No. 06-09. Toyonaka: Graduate School of Economics and Osaka School of International Public Policy.
- Nishina, Kazuhiko, Nabil Maghrebi, and Mark J. Holmes. 2012. Nonlinear Adjustments of Volatility Expectations to Forecast Errors: Evidence from Markov-Regime Switches in Implied Volatility. *Review of Pacific Basin Financial Markets and Policies* 15: 1250007. [CrossRef]
- Pesaran, M. Hashiem, Yongcheol Shin, and Richard J. Smith. 2001. Bounds Testing Approaches to the Analysis of Level Relationships. *Journal of Applied Econometrics* 16: 289–326. [CrossRef]
- Pindyck, Robert S. 2001. The dynamics of commodity spot and futures markets: A Primer. *The Energy Journal* 22: 1–29. [CrossRef]
- Pouzo, Demian, Zacharias Psaradakis, and Martin Sola. 2022. Maximum likelihood estimation in Markov regime-switching models with covariate-dependent transition probabilities. *Econometrica* 90: 1681–710. [CrossRef]
- Raza, Naveed, Syed Jawad Hussain Shahzad, Aviral Kumer Tiwari, and Muhammad Shahbaz. 2016. Asymmetric impact of gold, oil prices and their volatilities on stock prices of emerging markets. *Resources Policy* 49: 290–301. [CrossRef]

- Sarwar, Ghulam. 2012. Is VIX an investor fear gauge in BRIC equity markets? *Journal of Multinational Financial Management* 22: 55–65. [CrossRef]
- Silvennoinen, Annastiina, and Susan Thorp. 2013. Financialization, crisis and commodity correlation dynamics. *Journal of International Financial Markets, Institutions and Money* 24: 42–65. [CrossRef]
- Smales, Lee A. 2022. Spreading the fear: The central role of CBOE VIX in global stock market uncertainty. *Global Finance Journal* 51: 100679. [CrossRef]
- Szado, Edward. 2009. VIX Futures and options: A Case Study of Portfolio Diversification During the 2008 Financial Crisis. *The Journal of Alternative Investments* 12: 68–85. [CrossRef]
- Tang, Ke, and Wei Xiong. 2012. Index Investment and the Financialization of Commodities. *Financial Analysts Journal* 68: 54–74. [CrossRef]
- Wang, Gang-Jin, and Chi Xie. 2012. Cross-correlations between WTI crude oil market and U.S. stock market: A perspective from econophysics. *Acta Physica Polonica B* 43: 10. [CrossRef]
- Whaley, Robert E. 2000. The investor fear gauge. *Journal of Portfolio Management* 26: 12–17. [CrossRef]
- Whaley, Robert E. 2009. Understanding the VIX. *The Journal of Portfolio Management* 35: 98–105. [CrossRef]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

Article

Causality between Arbitrage and Liquidity in Platinum Futures

Kentaro Iwatsubo ^{1,*}  and Clinton Watkins ^{1,2,*} 

¹ Graduate School of Economics, Kobe University, 2-1 Rokkodai-cho, Nada, Kobe 657-8501, Japan

² Faculty of International Liberal Arts, Akita International University, Yuwa, Akita-City 010-1292, Japan

* Correspondence: iwatsubo@econ.kobe-u.ac.jp (K.I.); clinton-watkins@aiu.ac.jp (C.W.)

Abstract: Arbitrage and liquidity are interrelated. Liquidity facilitates arbitrageurs' trading on deviations from the law of one price. However, whether arbitrage opportunity leads to an increase or decrease in liquidity depends on the cause of the deviation. A demand shock leads to greater liquidity, while asymmetric information is toxic to liquidity. We examine how arbitrage and liquidity influence each other in the world's largest platinum futures markets on exchanges in New York and Tokyo. The markets provide an interesting institutional setting because the futures are based on an identical underlying commodity but exhibit different liquidity characteristics both intraday and over their lifespans. Using intraday data, we find that deviation in currency-adjusted futures prices leads, on average, to an immediate increase in liquidity, suggesting that demand shocks are the dominant driver of arbitrage opportunities. Less actively traded futures experience a greater liquidity effect. Arbitrageurs improve liquidity in both New York and Tokyo by acting as discretionary liquidity traders and cross-sectional market-makers.

Keywords: arbitrage; efficiency; futures; liquidity; market integration; platinum

JEL Classification: G13; G14; G15; Q02



Citation: Iwatsubo, Kentaro, and Clinton Watkins. 2022. Causality between Arbitrage and Liquidity in Platinum Futures. *Journal of Risk and Financial Management* 15: 593. <https://doi.org/10.3390/jrfm15120593>

Academic Editor: Thanasis Stengos

Received: 3 November 2022

Accepted: 2 December 2022

Published: 9 December 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The law of one price (LOOP) suggests that the prices of futures on an identical underlying commodity that are traded on different exchanges should be the same, taking into account the currency of denomination and differences in contract specifications. Deviation in such prices gives rise to an arbitrage opportunity that, in competitive and efficient markets, would lead to prices equalizing across exchanges.

Liquidity has long been thought to influence the ability of market participants to conduct arbitrage in financial markets. High levels of market liquidity allow arbitrageurs to exploit price differences across markets that trade homogeneous, similar or closely linked securities. Low liquidity limits arbitrage, as traders who attempt to profit on the convergence of prices are exposed to higher transactions costs and a greater risk that prices move against their trades. Although in theory arbitrage refers to a riskless opportunity to profit from a pricing discrepancy, in practice, such opportunities are rare. Limits to arbitrage include fundamental risk, noise trader risk and implementation costs. Low liquidity or differences in liquidity between markets may enhance limits to arbitrage.

Recent research has examined how arbitrage may affect liquidity. Theory suggests that arbitrage profit opportunities may either lead to an improvement or a deterioration in liquidity, depending on the nature of the shock that led to the deviation in prices between markets. Arbitrage opportunities that arise due to non-fundamental demand shocks are liquidity-enhancing, as arbitrageurs enter both markets to transact against the direction of the shock (Gromb and Vayanos 2010; Holden 1995). Arbitrage trades are likely to be spread over time to minimise price impact, which would bring about autocorrelated orders and a persistent liquidity effect (Kyle 1985; Roll et al. 2007). Arbitrageurs trade against prevailing demand to provide liquidity to other market participants in exchange for a premium and,

in doing so, improve market integration (Rösch 2021). Under these circumstances, arbitrage profit opportunities would be expected to lead to an increase in liquidity.

Alternatively, arbitrage opportunities that arise due to information asymmetry may reduce liquidity due to adverse selection (Foucault et al. 2017; Kumar and Seppi 1994). So-called toxic arbitrage opportunities occur when new information is incorporated in one market's price, leading to a short-lived price deviation with respect to the other market (Johannsen 2017). Foucault et al.'s (2017) model suggests dealers respond to toxic arbitrage by widening bid-ask spreads to slow their trading and compensate for the risk of transacting at stale prices. Toxic arbitrage opportunities lead to a decrease in liquidity. Furthermore, limits to arbitrage may dissuade arbitrageurs from taking advantage of price deviations. For instance, arbitrageurs may avoid arbitrage trades with high idiosyncratic volatility that would expose them to losses or the need to liquidate the arbitrage trade (Shleifer and Vishny 1997).

These different strands of the literature raise several empirical questions. What is the direction of causality between arbitrage and liquidity? Are arbitrage opportunities generally associated with lower or higher liquidity? Does the sign and magnitude of the relationship between arbitrage and liquidity change when a futures contract is in a more or less actively traded phase of its lifespan?

Researchers have examined the empirical relationship between arbitrage and liquidity for similar or related securities in equity, currency and fixed-income markets. Roll et al. (2007) find bidirectional relationships between liquidity on the New York Stock Exchange (NYSE) and arbitrage opportunities in the futures cash basis associated with the NYSE composite index futures contract. Schultz and Shive (2010) show that arbitrage opportunities between dual-class shares, which are typically driven by the more liquid share, lead to increased trade volumes, with trades within one or the other share class being relatively more important in closing the price gap than matched trades across the dual-share classes. Marshall et al. (2013a) examines arbitrage opportunities in similar highly liquid exchange traded funds (ETFs) that track the S&P500 index. Their results suggest that a fall in liquidity combined with an increase in liquidity risk contribute to arbitrage opportunities, which are rapidly eliminated with buyer (seller)-initiated trades in the underpriced (overpriced) ETF that is most prevalent. Ghadhab (2018) find that arbitrage opportunities in cross-listed stocks lead to greater liquidity. Rösch (2021) examine the arbitrage and liquidity relationships between the American Depositary Receipt (ADR) market and home market shares and find that a positive shock to arbitrage decreased the price deviations and bid-ask spreads. Rappoport and Tuzun (2020) find bidirectional Granger causality between liquidity and arbitrage opportunity between ETFs and their constituents in both equity and bond markets, but the effect of arbitrage on liquidity and vice versa is larger for bond than equity markets, as bond ETF constituents are generally less actively traded. Foucault et al. (2017) examine triangular arbitrage in foreign exchange markets and find a positive relationship between illiquidity and both the fraction of toxic arbitrage opportunities and arbitrageurs' relative speed in trading.

Several recently published articles analyze commodity futures, including the platinum contract traded in New York. Ludwig (2019) investigate the relationship between speculative activity and liquidity, while Bohl et al. (2021) consider speculative activity and informational efficiency. Lauter and Prokopczuk (2022) identify low-frequency proxies for commodity futures market quality. Boos and Grob (2022) investigate the trading strategies of managed futures funds, Sakkas and Tessaromatis (2020) evaluate factor-based commodity futures investment, and Kwon et al. (2020) evaluate model momentum. Sun et al. (2023) examine the price impact of traders on commodity futures. Tokyo platinum contracts also feature in recent published research. Boubaker et al. (2021) evaluate the performance of momentum based on functional data analysis. Iwatsubo and Watkins (2020) present evidence on which traders influence the efficient prices of commodity futures traded in Tokyo. Iwatsubo et al. (2018) examine the intraday seasonality of the microstructure characteristics in the platinum and gold futures contracts traded in New York and Tokyo.

To our knowledge, there is no published research on the potentially bidirectional causal relationship between arbitrage and liquidity in commodity futures markets. However, the arbitrage–liquidity relationship is particularly relevant for commodity futures, given their institutional environment. It is now common for multiple futures exchanges to list contracts based on the same underlying commodity traded in concurrent or overlapping sessions in different countries. Exchanges use day and night trading sessions to facilitate market access by participants in different time zones around the world, arguably to boost liquidity and promote efficient price discovery. Arbitrageurs can trade on LOOP violations between exchanges around the clock. With sufficient liquidity, arbitrage activity between the exchanges is expected to encourage a single world price for futures based on an identical underlying commodity, consistent with the LOOP after adjusting for contract specifications and exchange rates. Hence, understanding the potentially bidirectional relationship between arbitrage and liquidity, and particularly whether arbitrage is beneficial or harmful for liquidity, is of practical importance in commodity futures markets. Our research contributes toward filling this gap in the literature.

In this paper, we examine the relationship between arbitrage and liquidity in the platinum futures markets on the New York Mercantile Exchange (NYMEX) and the Tokyo Commodity Exchange (TOCOM), taking into account both potential directions of causality. We use intraday price and volume data for one contract on each exchange with a common expiry month. The platinum futures markets in New York and Tokyo provide an interesting environment to examine the interaction of arbitrage and liquidity. The exchanges are the primary global derivative venues for hedging and speculation in platinum. Both exchanges trade futures based on the same grade of the underlying commodity, while the contract specifications and currency of denomination differ.

The liquidity patterns for the contracts in New York and Tokyo are distinct both intraday and over their contract lifespans. This provides an interesting institutional setting to examine the relationship between arbitrage and liquidity. Intraday liquidity on each market is greatest during the respective exchange’s daytime trading session (Iwatsubo et al. 2018). Prices may deviate from the parity implied by the LOOP due to the different intraday liquidity patterns if there are enhanced limits to arbitrage in the relatively illiquid market. As the relative liquidity conditions adjust over the trading day, arbitrageurs have the opportunity to exploit any deviation from the LOOP. Trading activity differs substantially over the two contracts’ lifespans. Near contracts are most actively traded in New York, while far contracts are most actively traded in Tokyo. In this paper, we focus on analyzing the differences in lifespan trading activity by dividing our data into subsamples that reflect the trading activity in New York and Tokyo.

Our analysis shows bidirectional Granger causality between arbitrage and liquidity. We find that deviation in currency-adjusted futures prices leads, on average, to an immediate increase in liquidity, suggesting that demand shocks are the dominant driver of arbitrage opportunities. The liquidity effect is relatively large when a contract is in a less actively traded phase of its lifespan compared with when the contract is more actively traded. A negative shock to liquidity in one market leads to room for arbitrage between the markets. Liquidity in the other market responds in the same direction, consistent with liquidity commonality.

Our research contributes to the small-but-growing literature that analyzes the bidirectional relationship between arbitrage and liquidity for similar securities that trade on different exchanges. Our paper extends this research to include commodity futures, an asset class in which the situation of multiple exchanges in different countries listing contracts based on an identical underlying commodity with overlapping trading sessions is common.

This article proceeds as follows. We explain the institutional details of the platinum markets, the data and variable construction in Section 2. We describe our empirical methodology in Section 3. In Section 4, we discuss our empirical results and their interpretation. Section 5 concludes the paper.

2. Platinum Futures Market Details and Data

2.1. Market Institutional Details

The platinum futures markets on NYMEX and TOCOM are the world's largest.¹ During the period of this study, NYMEX was the bigger market both in terms of the weight of metal represented by the futures traded as well as the number of contracts traded. However, TOCOM has long been an important market for hedging and speculating on platinum in futures, and it was by far the largest market in the early 2000s. Both exchanges are major venues for hedging by firms involved in producing or consuming platinum. Automobile catalytic converters constitute the largest use of platinum at around 40% of global demand, followed by jewelery at 30% (Johnson Matthey 2016; McDonald and Hunt 1982). Platinum is also used in the chemical, electronics, glass, medical and petroleum industries. South Africa is the largest producer of platinum, constituting over 70% of the global supply, while Russia produces around 12% (Johnson Matthey 2016). Platinum derivatives have long been used by professional investors to take exposure to the market, and the introduction of platinum ETFs since 2010 has provided access at a lower cost, particularly for smaller investors (Vigne et al. 2017).

The physical commodities underlying the contracts on NYMEX and TOCOM are identical at a minimum of 99.95 percent pure platinum, but the contract units differ in terms of the weight of the metal. The NYMEX contract is based on 50 troy ounces, while the TOCOM standard contract is for 500 g, equivalent to 16.08 troy ounces.²

The exchanges have different contract listing schedules. NYMEX lists monthly contracts for the nearest three consecutive months and contracts expiring in January, April, July and October for the nearest 15 months. TOCOM lists the nearest six contract months with expiries in February, April, June, August, October and December. The most actively traded contracts also differ between the exchanges. Contracts for the nearest four months to expiry after the spot month are the most actively traded on NYMEX, while the farthest and second-farthest listed contracts are most active on TOCOM. Position limit rules constrain trading in the spot or expiry month on both exchanges.³

Figure 1 shows the daily prices for the April 2016 expiration of New York and Tokyo platinum futures in the units in which each contract is denominated. The contracts listed on NYMEX are in US dollars per troy ounce, while those on TOCOM are in Japanese yen per gram. Platinum prices trend down from May 2015 to January 2016, with industry analysts observing factors including a declining production deficit, greater investment demand offset by higher South African mine production and lower jewelery demand. From February to May 2016, prices increase with market expectations of a larger production deficit in 2016 on lower South African mining output, greater demand for autocatalysts in Europe, a recovery in jewelery demand and continued investment demand from Japan (Johnson Matthey 2016; Wilson 2015).

2.2. Data and Sample

Our raw data consists of 1 min best bid and ask quotes and traded volumes for the April 2016 expiration platinum futures contracts traded on NYMEX and TOCOM. Data for the NYMEX contract were obtained from Thomson Reuters (now Refinitiv), and those for the TOCOM contract were obtained directly from the exchange. The data begin at 9:00 a.m. Japan Standard Time (JST) on 7 May 2015 and end at 4:00 a.m. JST on 23 April 2016, containing 264 trading days and 223,650 1 min observations.

At the time of the April 2016 contracts, the TOCOM daytime trading session was scheduled to begin at 9:00 a.m. JST and end at 3:15 p.m. JST. After a break, the night session began at 4:30 p.m. JST and ended at 4:00 a.m. JST the next morning.⁴ For the purpose of this study, we define the trading day to be 19 continuous hours long based on the TOCOM day and night trading sessions. This covers the daily period for which the NYMEX and TOCOM platinum trading sessions overlap for our sample period.

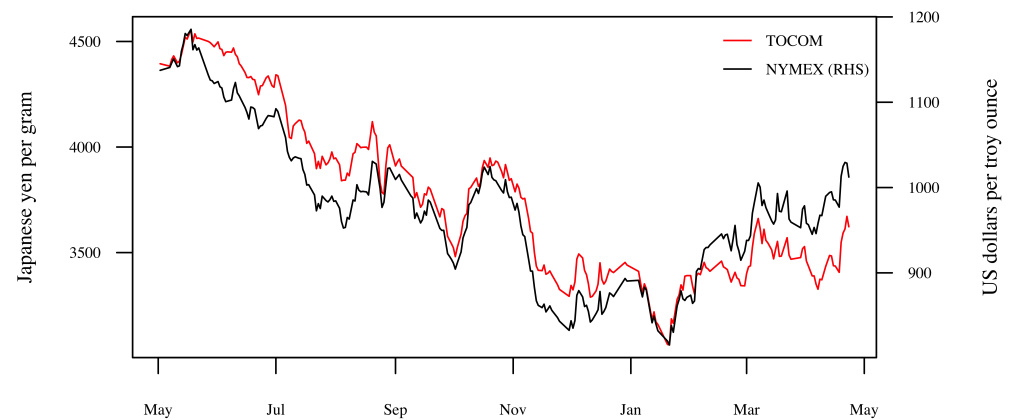


Figure 1. Daily prices for the April 2016 futures on TOCOM and NYMEX. (Note: The figure shows the daily April 2016 futures prices for platinum on NYMEX and TOCOM in US dollars per troy ounce on the RHS axis and Japanese yen per kilogram on the LHS axis).

We convert the NYMEX prices from US dollars per troy ounce to yen per gram using the troy ounce to gram conversion factor and 1 min USD/JPY forward exchange rates calculated from the 1 min spot USD/JPY exchange rate and daily US dollar and Japanese yen LIBOR rates.⁵ The foreign exchange and interest rate data were obtained from Bloomberg.

To take advantage of the differences between the trading activity over the lifespan of the New York and Tokyo contracts, we divide our data into five subsamples as described in Table 1. Subsample 1, running from 7 May to 2 July 2015, reflects the period in which the Tokyo future is the farthest contract and the most traded TOCOM contract. Subsample 2 covers the period 6 July to 4 September 2015, when the Tokyo future is the second-farthest contract. Trade in the New York future is relatively low but increasing over time during subsamples 1 and 2. Subsample 3, from 9 September to 30 November 2015, reflects the period in which there is consistent trade in both the New York and Tokyo contracts, but neither is the most traded contract on its respective exchange. Subsample 4, from 1 December 2015 to 31 March 2016, covers the period in which the New York contract is most actively traded (i.e., when the NYMEX contract was within the nearest four months to expiry). During this period, the Tokyo contract continues to be traded, albeit at a relatively low level compared with subsample 1. Subsample 5 represents the expiration month for both contracts and runs from 1 to 23 April 2016. We omit this period from our modeling because the exchanges impose position limits that may have influenced our results.

Table 1. The April 2016 futures contracts on NYMEX and TOCOM and the subsamples.

Exchange	2015						2016					
	May	Jun.	Jul.	Aug.	Sep.	Oct.	Nov.	Dec.	Jan.	Feb.	Mar.	Apr.
TOCOM	Farthest		2nd Farthest		3rd Farthest		3rd Nearest		2nd Nearest		Nearest	
NYMEX	11	10	9	8	7	6	5	4	3	2	1	Expiry
Subsamples	1		2		3		4		5			

Notes: The TOCOM contract is denoted using the nearest and farthest nomenclature. The NYMEX contract is labeled in terms of months before the expiration month. The shaded regions indicate the five subsamples.

2.3. Variable Construction

To analyze the relationships between arbitrage and liquidity across the platinum futures markets in Tokyo and New York, measures of the arbitrage profit opportunity and liquidity are required. We calculate arbitrage profit (PROF) at time t , $PROF_t$, based on an arbitrage transaction between the April 2016 platinum futures on NYMEX and TOCOM in Equation (1). PROF is defined as the supremum of buying the NYMEX contract and selling

the TOCOM contract, selling the NYMEX contract and buying the TOCOM contract or zero in the case that neither trade provides a positive arbitrage profit. PROF is expressed relative to Tokyo’s mid-price. This profit measure has the advantage of being a directly tradeable arbitrage strategy. PROF can only take positive or zero values. A positive value for PROF indicates the violation of the LOOP and a possible arbitrage opportunity if the deviation in price is enough to offset the costs and risk of transacting. A PROF of zero indicates that no arbitrage opportunity exists. PROF is defined as

$$PROF_t = \sup \left[\frac{Bid_t^{TO} - Ask_t^{NY} \times FX_t}{Mid_t^{TO}}, \frac{Bid_t^{NY} \times FX_t - Ask_t^{TO}}{Mid_t^{TO}}, 0 \right] \quad (1)$$

where Bid_t^i and Ask_t^i are the bid and ask quotes and Mid_t^i is the mid-prices at time t on platinum market $i = \{NY, TO\}$, representing NYMEX and TOCOM, respectively. The mid-prices are calculated as $Mid_t^i = (Bid_t^i + Ask_t^i)/2$. FX_t refers to the USD/JPY forward exchange rate at time t .

The relative quoted bid–ask spread is used to measure the market liquidity of each exchange. The higher the relative quoted spread, the less liquid the market is. The spread SP_t^i is defined as

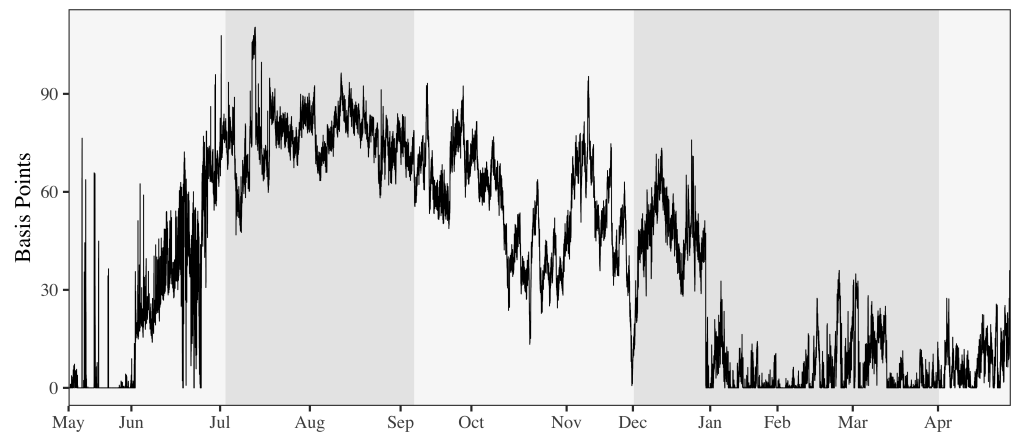
$$SP_t^i = \left(\frac{Ask_t^i - Bid_t^i}{Mid_t^i} \right) \quad (2)$$

In the modeling that follows, we denote the spread on NYMEX as SPNY and that on TOCOM as SPTO. To reduce the impact of microstructure noise, we convert the data to a 5 min frequency by taking the mean of PROF, SPNY and SPTO for each 5 min interval. This leaves us with 41,322 observations for the full sample. Figure 2 shows the 5 min data for PROF, SPTO and SPNY in panels (a), (b) and (c), respectively. The shaded regions indicate the five subsamples.

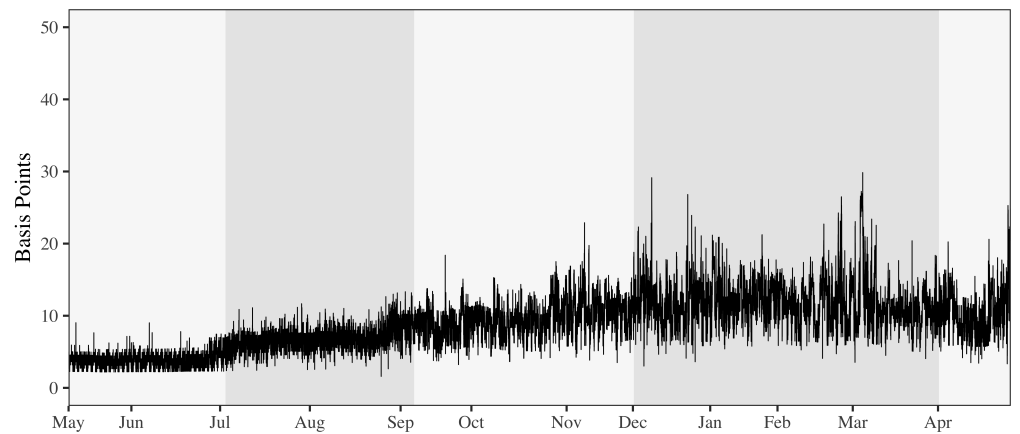
Table 2 shows the summary statistics for PROF, SPNY and SPTO for the full sample and each subsample. PROF is relatively high in subsample 2 followed by subsample 3 when the Tokyo contract is not the farthest and the New York contract is not within the nearest 4 months to expiration. PROF is lower and less variable in subsample 4 when the New York contract is actively traded compared with subsample 1, when the Tokyo contract is actively traded. The bid–ask spread on TOCOM increases and is more variable over successive subsamples, while the spread on NYMEX declines and becomes less variable, as would be expected given that the far contract is the most traded in Tokyo, while the near-expiration contracts are the most traded in New York. All variables are stationary for the full sample and in all subsamples. As the subsamples are formed on the basis of when the contracts are most actively traded in each market, the number of observations differs.

Figure 3 shows the number of platinum futures traded during each 5 min period over the full sample for TOCOM (VOLTO) and NYMEX (VOLNY) in panels (a) and (b), respectively. Panel (c) provides a cumulative sum of the trading volume in the April 2016 contract on each exchange. This shows the different pattern in trading volume over the lifespan of the contracts on each exchange. The TOCOM April 2016 contract is most actively traded when it is the far contract, while the NYMEX contract is most active when it is within the nearest four contracts.

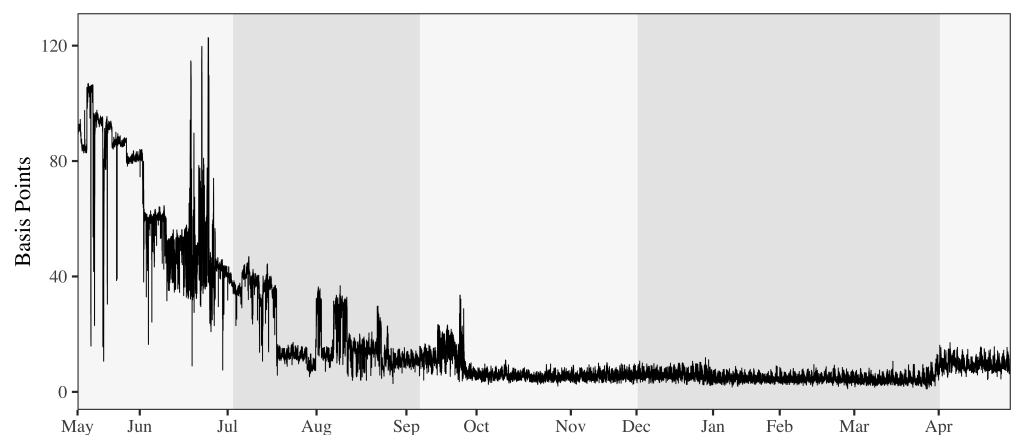
Table 3 provides a summary of the statistics for the 5 min aggregate trade volumes on each exchange. The different trading activity patterns between New York and Tokyo over the respective contract lifespans can be observed in the mean and maximum number of contracts traded in each 5 min interval and in the sum of all contracts traded over each subsample. The mean, maximum and standard deviation of volumes decreases over subsamples 1–4 for TOCOM and increases for NYMEX. The minimum of the 5 min volumes is zero for each subsample on both exchanges (not shown in the table). The volume data series are highly leptokurtic.



(a) Arbitrage profit (PROF)



(b) TOCOM bid-ask spread (SPTO)



(c) NYMEX bid-ask spread (SPNY)

Figure 2. Arbitrage profit and liquidity variables. (Note: The figure shows the 5 min of average arbitrage profit (PROF) and the relative bid-ask spreads for the NYMEX (SPNY) and TOCOM (SPTO) contracts calculated from 1 min data for 264 trading days starting at 9:00 a.m. Japan Standard Time (JST) on 7 May 2015 and ending at 4:00 a.m. JST on 23 April 2016 in basis points in panels (a–c), respectively. The shaded regions indicate the five subsamples).

Table 2. Summary statistics for the arbitrage and liquidity variables.

	Full Sample			Subsample 1			Subsample 2		
	PROF	SPTO	SPNY	PROF	SPTO	SPNY	PROF	SPTO	SPNY
Mean	0.0040	0.0009	0.0020	0.0024	0.0004	0.0066	0.0077	0.0007	0.0020
Median	0.0043	0.0009	0.0007	0.0021	0.0004	0.0061	0.0077	0.0007	0.0014
Minimum	0.0000	0.0002	0.0001	0.0000	0.0002	0.0008	0.0047	0.0002	0.0003
Maximum	0.0110	0.0030	0.0127	0.0108	0.0009	0.0127	0.0110	0.0013	0.0047
St. Dev.	0.0031	0.0004	0.0025	0.0025	0.0001	0.0021	0.0009	0.0001	0.0011
Skewness	−0.03	0.56	1.83	0.66	0.38	0.12	0.18	0.53	0.75
Ex. Kurtosis	−1.46	0.81	2.32	−0.73	0.84	−1.19	1.75	1.51	−1.00
Unit Root	−5.66	−61.13	−10.09	−12.76	−54.53	−16.92	−7.23	−58.49	−8.05
P-value	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
Observations	41,322	41,322	41,322	7455	7455	7455	8946	8946	8946

	Subsample 3			Subsample 4			Subsample 5		
	PROF	SPTO	SPNY	PROF	SPTO	SPNY	PROF	SPTO	SPNY
Mean	0.0055	0.0010	0.0007	0.0015	0.0012	0.0005	0.0007	0.0011	0.0010
Median	0.0057	0.0010	0.0006	0.0005	0.0012	0.0005	0.0006	0.0010	0.0010
Minimum	0.0001	0.0003	0.0002	0.0000	0.0003	0.0001	0.0000	0.0003	0.0004
Maximum	0.0095	0.0023	0.0034	0.0076	0.0030	0.0013	0.0036	0.0025	0.0017
St. Dev.	0.0016	0.0002	0.0004	0.0019	0.0003	0.0001	0.0006	0.0003	0.0002
Skewness	−0.33	0.49	2.01	1.17	1.22	0.83	0.68	1.31	0.55
Ex. Kurtosis	−0.29	0.63	4.02	−0.09	3.51	1.04	−0.28	2.64	0.37
Unit Root	−3.91	−34.24	−22.98	−5.12	−30.88	−60.67	−7.24	−13.37	−24.26
P-value	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
Observations	10,437	10,437	10,437	14,484	14,484	14,484	3408	3408	3408

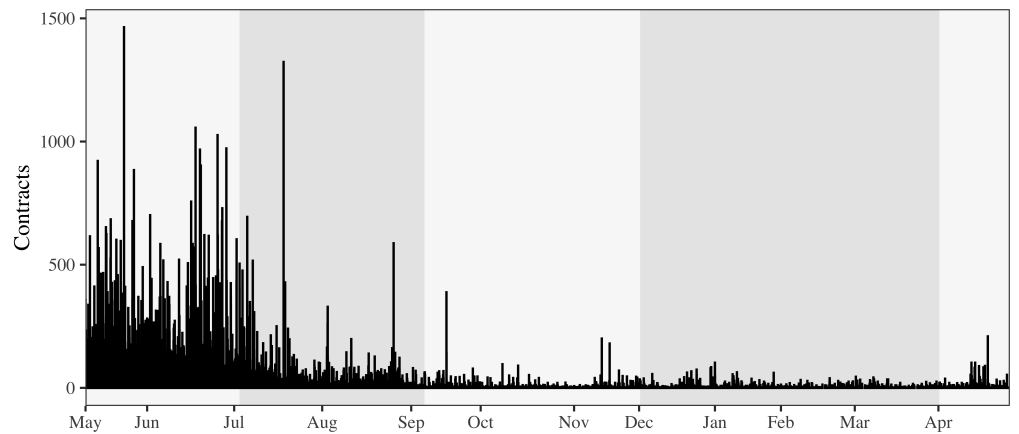
Notes: PROF represents arbitrage profit. SPNY and SPTO are the relative bid–ask spreads for the April 2016 platinum futures contracts traded on NYMEX and TOCOM, respectively. Each variable is constructed using 1 min data, and the result is averaged over 5 min periods. The table shows summary statistics for the full sample and each subsample calculated from the 5 min data. Unit Root and P-value represent the Phillips–Perron test statistic and its associated p-value for a 5% level of significance.

Table 3. Summary statistics for trading volume during each 5 min interval for each sample.

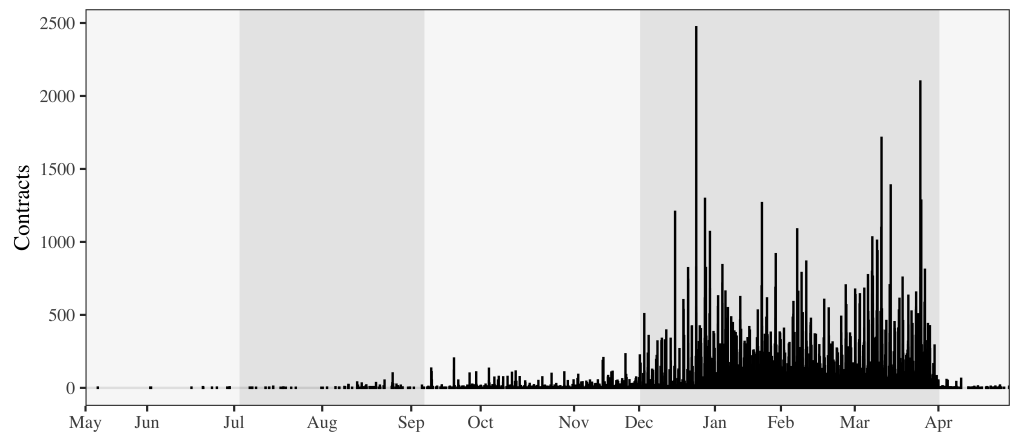
	Full Sample		Subsample 1		Subsample 2	
	VOLTO	VOLNY	VOLTO	VOLNY	VOLTO	VOLNY
Mean	11.08	16.75	46.02	0.00	10.51	0.10
St. Dev.	39.65	58.04	77.36	0.10	29.96	1.59
Skewness	10.90	10.32	5.57	31.34	15.89	36.71
Ex. Kurtosis	202.98	215.55	50.81	1150.56	501.97	1885.75
Maximum	1464	2471	1464	5	1323	99
Sum	457,723	692,167	343,072	30	94,006	904

	Subsample 3		Subsample 4		Subsample 5	
	VOLTO	VOLNY	VOLTO	VOLNY	VOLTO	VOLNY
Mean	1.03	1.33	0.69	46.77	1.56	0.19
St. Dev.	6.48	7.83	3.47	90.43	6.98	2.07
Skewness	29.02	15.88	11.36	6.76	13.71	21.45
Ex. Kurtosis	1403.95	340.72	189.45	93.15	292.49	526.09
Maximum	388	230	102	2471	209	62
Sum	10,713	13,833	9932	677,400	5321	632

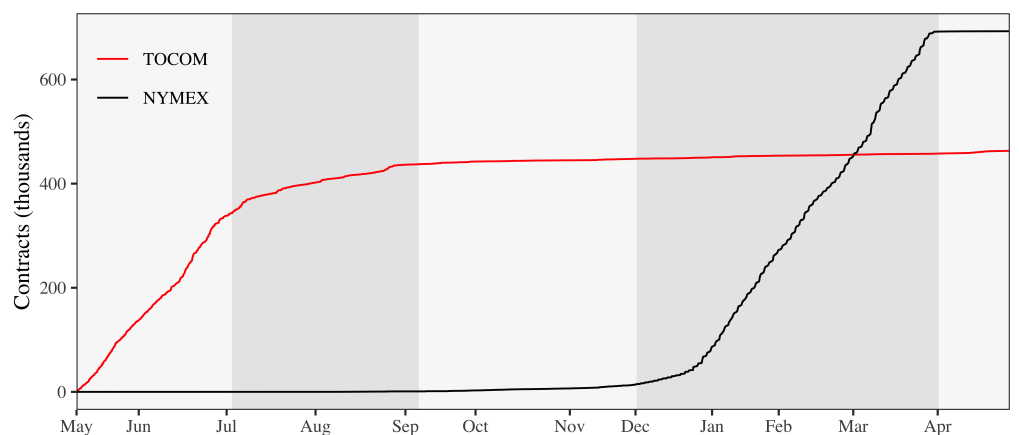
Notes: VOLTO and VOLNY represent the aggregate 5 min trading volumes in the April 2016 platinum futures contracts traded on TOCOM and NYMEX, respectively. Maximum indicates the largest number of contracts traded within the 5 min time intervals. The minimum for all subsamples is zero contracts for both exchanges and is not shown in the table. Sum provides the total number of contracts traded during each subsample for each exchange.



(a) TOCOM trading volume (VOLTO)



(b) NYMEX trading volume (VOLNY)



(c) Cumulative volume on each exchange

Figure 3. Platinum futures contract volumes on TOCOM and NYMEX (Notes: Panels (a,b) show the 5-min aggregate trading volumes for the TOCOM (VOLTO) and NYMEX (VOLNY) April 2016 platinum futures contracts over 264 trading days starting at 9:00 a.m. Japan Standard Time (JST) on 7 May 2015 and ending at 4:00 a.m. JST on 23 April 2016, respectively. Panel (c) provides the cumulative total trading volume for the April 2016 contract on each exchange over the sample period. The shaded regions indicate the five subsamples).

3. Methodology

We estimated a vector autoregression (VAR) of three variables—SPNY, SPTO and PROF—as follows:

$$y_t = A_1y_{t-1} + \dots + A_p y_{t-p} + CD + e_t \tag{3}$$

where $y_t = [SPNY_t, SPTO_t, PROF_t]'$, a constant and deterministic trend are included with $D = [CONST, TREND]'$, A represents the coefficients of the lagged variables and C contains the coefficients of the constant and deterministic trend terms.

The model is estimated efficiently using ordinary least squares for each of subsamples 1–4. We select the number lags to include in each model using the Schwarz Information Criterion. The number of lags included in the models for subsamples 1–4 are 6, 5, 9 and 8, respectively. We order the variables as SPNY, SPTO and PROF. The ordering of the variables in the VAR can influence the orthogonalized impulse responses, and if this happens, the variables should be ordered from the most to least exogenous. However, given that we view both arbitrage and liquidity as endogenous, it is difficult to determine an appropriate variable ordering. Accordingly, we calculate generalized impulse response functions (GIRFs) using the method suggested by Pesaran and Shin (1998), which are not influenced by the order of the variables in the VAR.

4. Empirical Results

Granger causality tests based on the estimated VAR models suggest a bidirectional relationship between the arbitrage profits and liquidity (see Table 4). The null of no Granger causality is rejected at the 1-percent level for all tests, except PROF does not Granger cause SPNY or SPTO in subsample 2, being rejected at the 5-percent level, and SPNY does not Granger cause SPTO or PROF, being rejected at the 10-percent level.

Table 4. Granger causality F-statistics.

Null Hypothesis	Subsample 1	Subsample 2	Subsample 3	Subsample 4
PROF does not Granger cause SPNY or SPTO	5.40 ***	2.12 **	2.32 ***	3.70 ***
SPTO does not Granger cause SPNY or PROF	5.55 ***	16.89 ***	6.73 ***	6.21 ***
SPNY does not Granger cause SPTO or PROF	6.48 ***	1.63 *	4.14 ***	11.63 ***

Notes: PROF represents arbitrage profit. SPNY and SPTO are the relative bid–ask spreads for the April 2016 platinum futures contracts traded on NYMEX and TOCOM, respectively. The table shows the Granger causality F-statistic. ***, ** and * indicate significance at the 1%, 5% and 10% levels, respectively.

Plots of the GIRFs derived from the VAR model are presented in Figures 4–9, organized as single impulse response pairs shown for each of subsamples 1–4 in panels (a–d), respectively. The GIRFs are shown for 200 steps ahead with 90-percent confidence intervals bootstrapped using 200 iterations. Note that 50 steps ahead represents just over 4 hours of trading time.

Figures 4 and 5 indicate that a one standard deviation shock in PROF is associated with an immediate and generally short-lived negative response in the relative bid–ask spreads of both New York and Tokyo. This suggests that, in the short term, liquidity improves in both markets following an arbitrage profit opportunity, as arbitrageurs act as “cross-sectional market makers” as described by Holden (1995). Accordingly, we conclude that the deviations from the LOOP that lead to arbitrage opportunities in platinum futures are generally due to non-fundamental demand shocks.

Furthermore, the immediate liquidity-enhancing effect of an arbitrage profit shock is greater when the futures contract is less actively traded. The size of the immediate liquidity response is greatest for SPNY in subsample 1, as shown in Figure 4a, which is when the New York contract is least actively traded. The opposite is true for SPTO in that the liquidity response is greatest in subsample 4, as shown in Figure 5d, at the time when the Tokyo contract is least actively traded. The peak of the immediate liquidity response in New York decreases monotonically over subsamples 1–4, while it increases in Tokyo. The

liquidity-enhancing effect of arbitrage profit opportunities is not only greater but also more persistent when a market is less actively traded, consistent with the findings of Rappoport and Tuzun (2020) for equity and bond markets.

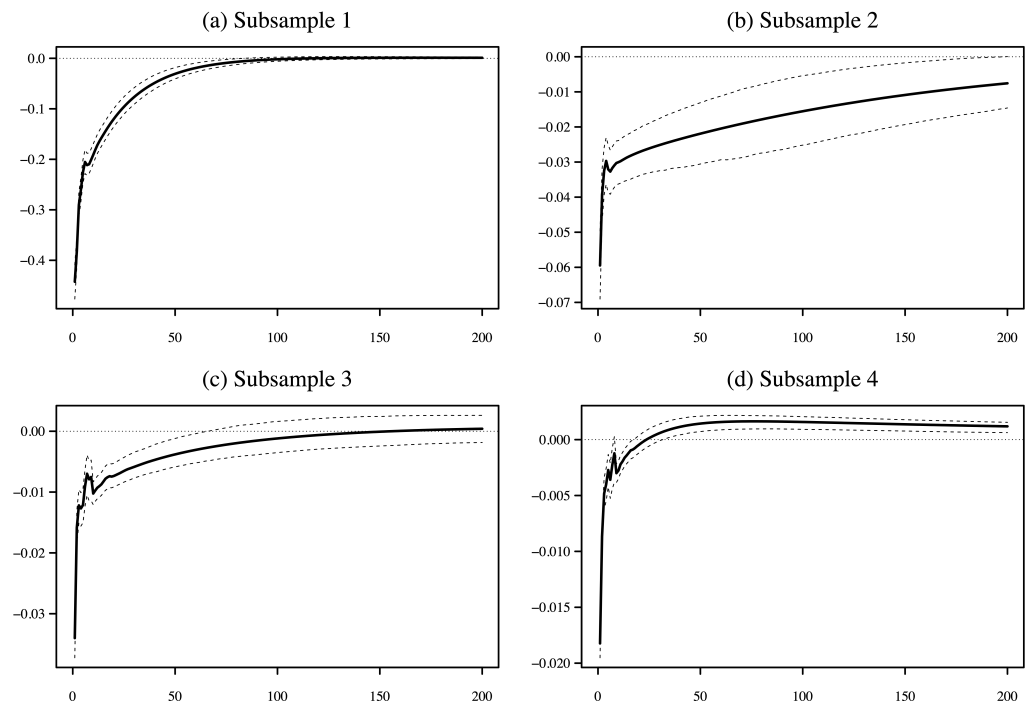


Figure 4. Response of SPNY to a shock in PROF. (Notes: The solid lines represent the 200-step-ahead GIRFs. The dashed lines show 90-percent confidence intervals bootstrapped using 200 iterations).

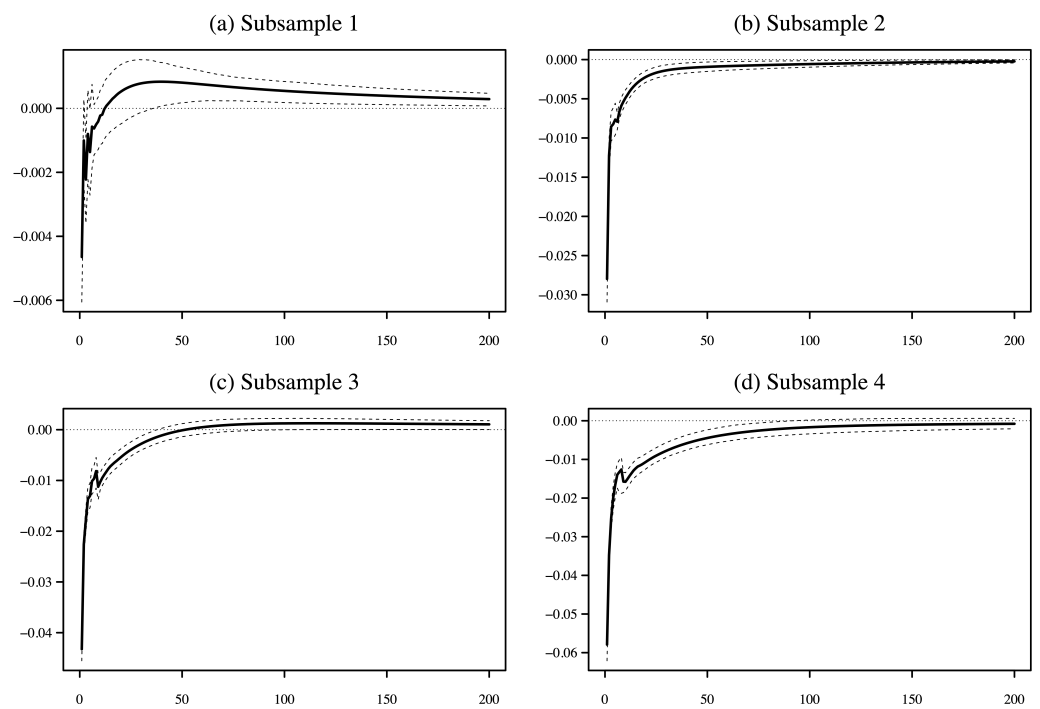


Figure 5. Response of SPTO to a shock in PROF. (Notes: The solid lines represent the 200-step-ahead GIRFs. The dashed lines show 90-percent confidence intervals bootstrapped using 200 iterations).

An interesting feature of the GIRFs when the contract in each market is most actively traded, which is in subsample 4 for New York shown in Figure 4d and in subsample 1 for Tokyo shown in Figure 5a, is the small but persistent positive marginal spreads that began within 50 steps following the PROF shock. This may seem inconsistent with our non-fundamental demand shock explanation. Does this delayed decrease in liquidity suggest the arbitrage opportunity is first mistaken for a non-fundamental demand shock and shortly after recognized to be the result of asymmetric information? Given that we observe this effect only when the respective contract is most actively traded, we favor an alternative and more plausible interpretation. In response to a demand shock, the resulting highly liquid market can lead to a deterioration in liquidity, as predicted by Admati and Pfleiderer (1988). Discretionary liquidity traders prefer to trade when they have little impact on prices to minimize their losses to informed traders. Greater liquidity trading encourages more informed trading at the same time, but competition between informed traders reduces their total profit. This benefits liquidity traders and encourages their further participation. Liquidity will eventually deteriorate when the concentration of informed traders increases the cost of adverse selection. Thus, for highly liquid markets, arbitrage associated with a non-fundamental demand shock may lead to an immediate increase in liquidity, followed by a persistent decrease in liquidity relative to before the shock.

Figures 6 and 7 show the response of PROF to a one standard deviation positive shock in SPNY and SPTO, respectively. The immediate response is negative in all cases, which seems inconsistent with the notion that a decrease in liquidity should increase arbitrage profit opportunities. However, the immediate negative response is explained by how the widening of the relative spread affects the calculation of arbitrage profit. A positive shock in the relative spread represents a decrease in or negative shock to liquidity and occurs via an increase in the ask price, a decrease in the bid price or both. The change in ask or bid prices due to the shock decreases PROF on impact, as Equation (1) suggests. As time passes, the deviation from the LOOP increases as traders find more room for arbitrage following the negative liquidity shock. The positive response of PROF is most apparent for liquidity shocks originating in subsample 4 for New York (Figure 6d) and in subsample 1 for Tokyo (Figure 7a). Thus, the positive effect on arbitrage opportunity is larger when the negative liquidity shock comes from a market in the actively traded phase of its lifespan. The positive response is less apparent in Figure 6a–c for subsamples 1, 2 and 3, respectively, and Figure 7b–d for subsamples 2, 3 and 4, respectively. This is because the liquidity shock in these subsamples originates from a market with relatively few active traders and is less likely to lead to a meaningful increase in opportunities for arbitrage between New York and Tokyo. The magnitudes of the PROF responses to liquidity shocks from each spread are comparable, except in Figure 6a for subsample 1, when the New York market is relatively inactive.

Figures 8 and 9 show the response of SPTO to a shock in SPNY and the response of SPNY to a shock in SPTO, respectively. A positive shock in the relative spread of one market initially results in an increase in the relative spread of the other market. This reflects liquidity commonality, defined as the co-movement in liquidity across securities or markets that are driven by one or more common factors. Chordia et al. (2000) identified liquidity commonality in U.S. stocks, while Marshall et al. (2013b) and Caporin et al. (2015) provided evidence for a systematic liquidity factor in commodity futures markets. However, as time goes on, the positive response weakens, and spreads generally narrow as traders find that the source of the spread increase is non-fundamental. The latter liquidity-enhancing effect dominates in all subsamples, except for subsample 1 in Figure 8a. The liquidity-enhancing effect is of greater magnitude when the impulse market is in a more active part of its lifespan and the response market is less active. This can be observed clearly in Figure 8d for subsample 4 and Figure 9a for subsample 1.

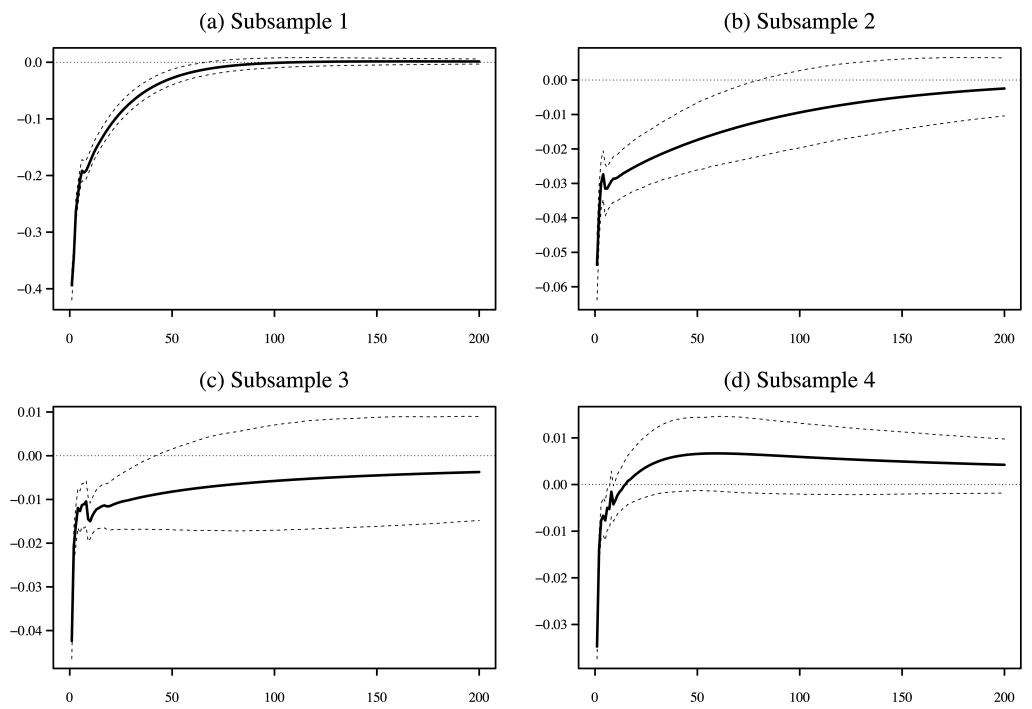


Figure 6. Response of PROF to a shock in SPNY. (Notes: The solid lines represent the 200-step-ahead GIRFs. The dashed lines show 90-percent confidence intervals bootstrapped using 200 iterations).

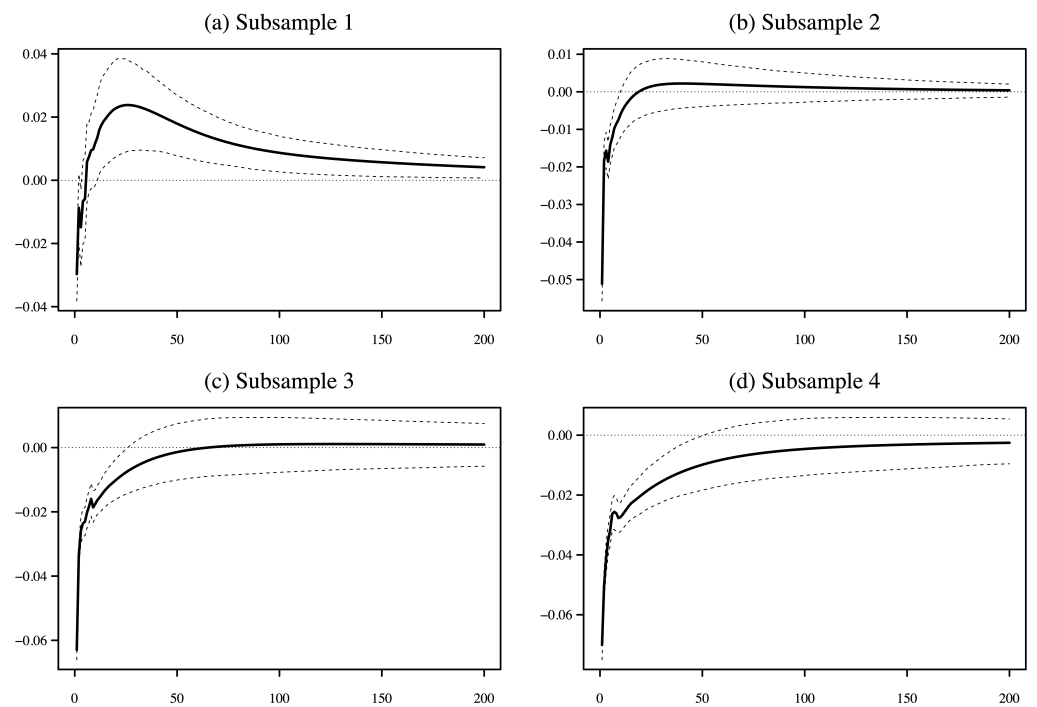


Figure 7. Response of PROF to a shock in SPPT. (Notes: The solid lines represent the 200-step-ahead GIRFs. The dashed lines show 90-percent confidence intervals bootstrapped using 200 iterations).

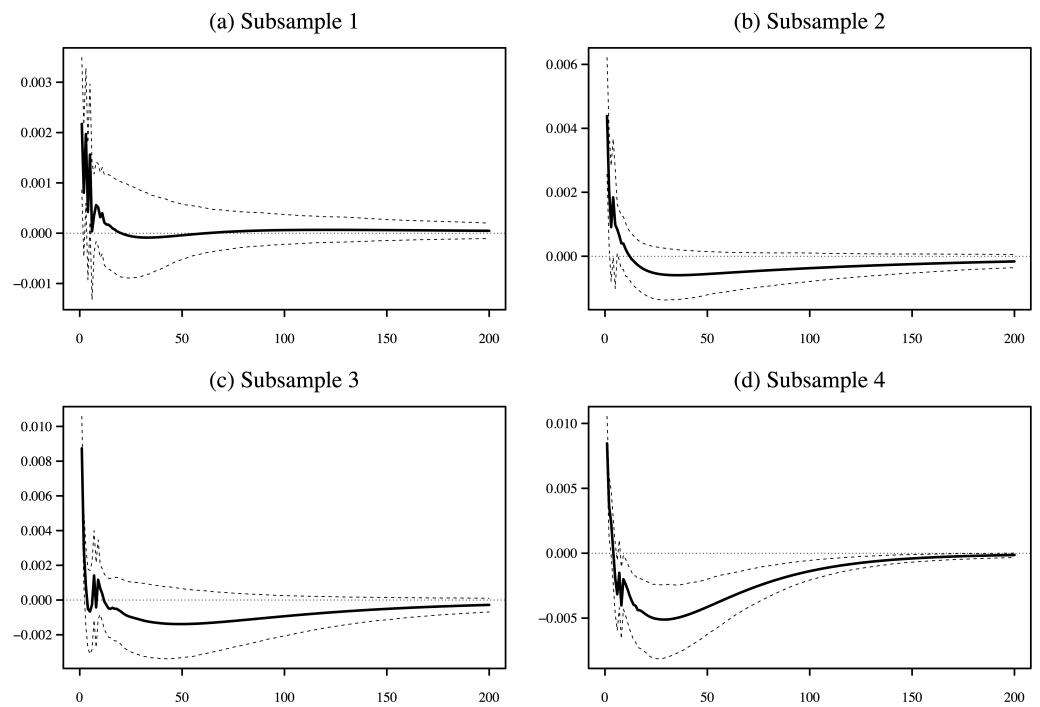


Figure 8. Response of SPTO to a shock in SPNY. (Notes: The solid lines represent the 200-step-ahead GIRFs. The dashed lines show 90-percent confidence intervals bootstrapped using 200 iterations).

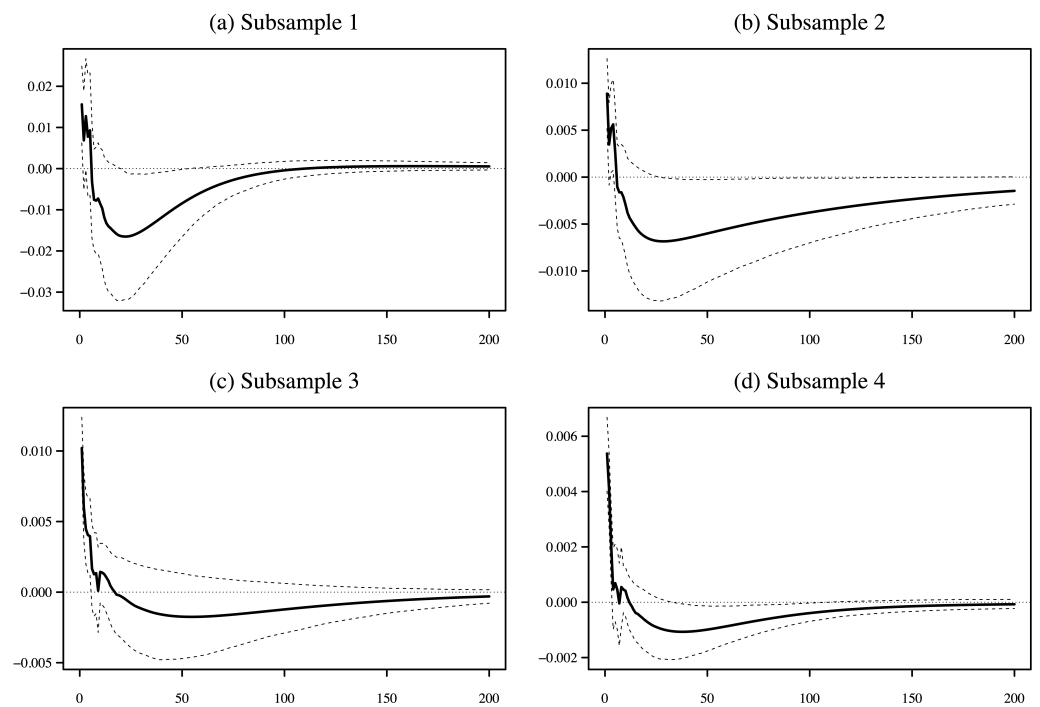


Figure 9. Response of SPNY to a shock in SPTO. (Notes: The solid lines represent the 200-step-ahead GIRFs. The dashed lines show 90-percent confidence intervals bootstrapped using 200 iterations).

5. Conclusions

Arbitrage and liquidity in markets for substitute securities are closely related. Theory posits reasons by which arbitrage should affect liquidity and liquidity should affect arbitrage. Whether arbitrage opportunities are associated with enhanced or diminished liquidity depends on the reason behind why the arbitrage opportunity has arisen. Non-

fundamental demand shocks enhance liquidity, while asymmetric information diminishes or is toxic to liquidity.

Arbitrage profit opportunities in futures markets trading an identical underlying grade of platinum in New York and Tokyo generally lead to increased liquidity in both markets. This suggests that non-fundamental demand shocks are the predominant cause of deviations from the LOOP across the two markets. The liquidity improvement is larger during periods of a contract's lifespan in which it is less actively traded. When a contract is most actively traded, liquidity first improves in response to an arbitrage opportunity and then deteriorates. This is consistent with the notion proposed by Admati and Pfleiderer (1988) that discretionary liquidity and informed traders may concentrate their transactions following non-fundamental demand shocks, and subsequently, liquidity declines as the risk of adverse selection rises. In particular, when the contracts are in an actively traded phase of their lifespan, a negative liquidity shock is associated with greater room for arbitrage between the two major future markets for platinum. A degree of relative liquidity commonality was observed across the two markets over the entirety of the contracts' lifespans.

Our results suggest that non-fundamental demand shocks, which are the main source of arbitrage opportunities between the New York and Tokyo markets, provide opportunities for both hedgers and speculators to trade, facilitated by liquidity provision from market participants performing the role of discretionary liquidity traders. Demand shocks provide a relatively advantageous opportunity for market participants to transact even when the contract in question is not in an actively traded phase of its lifespan.

Understanding the causal relationship between arbitrage and liquidity is important for commodity futures markets, given that multiple exchanges list contracts based on identical underlying commodities which are traded concurrently in overlapping day and night trading sessions. There is a number of potential directions for extending the research on this relationship in commodity futures markets. Although our research suggests toxic arbitrage is not prevalent in platinum futures, a switching regime approach may be used to better understand the circumstances around which arbitrage leads to lower market liquidity. Panel VAR models may be used to examine a series of contracts over time for a single underlying commodity or for multiple related commodity futures contracts at a point in time.

Author Contributions: Conceptualization, K.I. and C.W.; methodology, K.I. and C.W.; software, C.W.; validation, K.I. and C.W.; formal analysis, C.W.; investigation, K.I. and C.W.; resources, K.I. and C.W.; data curation, K.I.; writing—original draft preparation, C.W.; writing—review and editing, K.I. and C.W.; visualization, C.W.; supervision, K.I.; project administration, K.I.; funding acquisition, K.I. All authors have read and agreed to the published version of the manuscript.

Funding: This work is supported by JSPS KAKENHI Grant Numbers 19K01754 and 22K18529.

Data Availability Statement: Data were obtained from Thomson Reuters (now Refinitiv), Bloomberg and the Tokyo Commodity Exchange.

Acknowledgments: We thank Michael McAleer for his helpful comments and Tao Xu for assistance in compiling the data.

Conflicts of Interest: The authors declare no conflict of interest.

Notes

¹ NYMEX is part of the CME Group, and TOCOM is part of the Japan Exchange Group.

² The NYMEX platinum contract specifications can be found at <https://www.cmegroup.com/markets/metals/precious/platinum.contractSpecs.html> (accessed on 3 May 2019).

The TOCOM platinum standard contract specifications can be found at <https://www.jpx.co.jp/english/derivatives/products/precious-metals/platinum-standard-futures/01.html> (accessed on 3 May 2019).

³ Note that in addition to limits in the expiry month, TOCOM imposes looser position limits in the month before the expiry month and the second contract month. There is also a position limit for all contract months combined.

⁴ TOCOM extended its trading hours after the sample period of our study. The day session now opens at 8:45 a.m. JST and closes at 3:15 p.m. The night session runs from 4:30 p.m. to 6:00 a.m. the next day.

- ⁵ Forward exchange rates are used for the currency conversion presuming that the arbitrage trades between the New York and Tokyo markets are held to expiry. 12-month LIBOR rates are used when the futures have between 12 and 6 months to expiry, 6-month rates for when the contracts have between 6 and 3 months to expiry, the 2-month rates for when the contracts have between 2 months and 1 month to expiry, 1-month rates for when the contracts have between 1 month and 1 week to expiry, and 1-week rates are used for the contracts when they have less than 1 week to expiry. One gram is equivalent to 0.03215 troy ounces.

References

- Admati, Anat R., and Paul Pfleiderer. 1988. Theory of Intraday Patterns: Volume and Price Variability A Theory of Intraday Patterns: Volume and Price Variability. *The Review of Financial Studies* 1: 3–40. [CrossRef]
- Bohl, Martin T., Alexander Pütz, and Christoph Sulewski. 2021. Speculation and the informational efficiency of commodity futures markets. *Journal of Commodity Markets* 23: 100159. [CrossRef]
- Boos, Dominik, and Linus Grob. 2022. Tracking speculative trading. *Journal of Financial Markets* 100774. [CrossRef]
- Boubaker, Sabri, Zhenya Liu, Shanglin Lu, and Yifan Zhang. 2021. Trading signal, functional data analysis and time series momentum. *Finance Research Letters* 42: 101933. [CrossRef]
- Caporin, Massimiliano, Angelo Ranaldo, and Gabriel G Velo. 2015. Precious metals under the microscope: A high-frequency analysis. *Quantitative Finance* 15: 743–59. [CrossRef]
- Chordia, Tarun, Richard Roll, and Avanidhar Subrahmanyam. 2000. Commonality in liquidity. *Journal of Financial Economics* 56: 3–28. [CrossRef]
- Foucault, Thierry, Roman Kozhan, and Wing Wah Tham. 2017. Toxic arbitrage. *Review of Financial Studies* 30: 1053–94. [CrossRef]
- Ghadhab, Imen. 2018. Arbitrage opportunities and liquidity: An intraday event study on cross-listed stocks. *Journal of Multinational Financial Management* 46: 1–10. [CrossRef]
- Gromb, Denis, and Dimitri Vayanos. 2010. *Limits of Arbitrage: The State of the Theory*. National Bureau of Economic Research, Working Paper 15821. Cambridge: National Bureau of Economic Research. . [CrossRef]
- Holden, Craig W. 1995. Index arbitrage as cross-sectional market making. *Journal of Futures Markets* 15: 423–55. [CrossRef]
- Iwatsubo, Kentaro, and Clinton Watkins. 2020. Who influences the fundamental value of commodity futures in Japan? *International Review of Financial Analysis* 67: 1–15. [CrossRef]
- Iwatsubo, Kentaro, Clinton Watkins, and Tao Xu. 2018. Intraday seasonality in efficiency, liquidity, volatility and volume: Platinum and gold futures in Tokyo and New York. *Journal of Commodity Markets* 11: 59–71. [CrossRef]
- Johannsen, Kolja. 2017. Toxic Arbitrage and Price Discovery. Available online: <https://ssrn.com/abstract=2950161> (accessed on 6 March 2019). [CrossRef]
- Johnson Matthey. 2016. *PGM Market Report May 2016*. Johnson Matthey Technology Review, 60. London: Johnson Matthey. [CrossRef]
- Kumar, Praveen, and Duane J. Seppi. 1994. Information and Index Arbitrage. *The Journal of Business* 67: 481–509. [CrossRef]
- Kwon, Kyung Yoon, Jangkoo Kang, and Jaesun Yun. 2020. Weekly momentum in the commodity futures market. *Finance Research Letters* 35: 1–8. [CrossRef]
- Kyle, Albert S. 1985. Continuous Auctions and Insider Trading. *Econometrica* 53: 1315–35. [CrossRef]
- Lauter, Tobias, and Marcel Prokopczuk. 2022. Measuring commodity market quality. *Journal of Banking and Finance* 145: 106658. [CrossRef]
- Ludwig, Michael. 2019. Speculation and its impact on liquidity in commodity markets. *Resources Policy* 61: 532–47. [CrossRef]
- Marshall, Ben R., Nhut H. Nguyen, and Nuttawat Visaltanachoti. 2013a. ETF arbitrage: Intraday evidence. *Journal of Banking and Finance* 37: 3486–98. [CrossRef]
- Marshall, Ben R., Nhut H. Nguyen, and Nuttawat Visaltanachoti. 2013b. Liquidity commonality in commodities. *Journal of Banking and Finance* 37: 11–20. [CrossRef]
- McDonald, Donald, and Leslie B. Hunt. 1982. *A History of Platinum and Its Allied Metals*. London: Johnson Matthey.
- Pesaran, H. Hashem, and Yongcheol Shin. 1998. Generalized impulse response analysis in linear multivariate models. *Economics Letters* 58: 17–29. [CrossRef]
- Rappoport, David Elias, and Tugkan Tuzun. 2020. *Arbitrage and Liquidity: Evidence from a Panel of Exchange Traded Funds*. Finance and Economics Discussion Series 2020-097. Washington, DC: Board of Governors of the Federal Reserve System. [CrossRef]
- Roll, Richard, Eduardo Schwartz, and Avanidhar Subrahmanyam. 2007. Liquidity and the Law of One Price: The Case of the Futures-Cash Basis. *Journal of Finance* 62: 2201–34. [CrossRef]
- Rösch, Dominik. 2021. The impact of arbitrage on market liquidity. *Journal of Financial Economics* 142: 195–213. [CrossRef]
- Sakkas, Athanasios, and Nikolaos Tassaromatis. 2020. Factor based commodity investing. *Journal of Banking and Finance* 115: 105807. [CrossRef]
- Schultz, Paul, and Sophie Shive. 2010. Mispricing of dual-class shares: Profit opportunities, arbitrage, and trading. *Journal of Financial Economics* 98: 524–49. [CrossRef]
- Shleifer, Andrei, and Robert W. Vishny. 1997. The limits of arbitrage. *Journal of Finance* 52: 35–55. [CrossRef]
- Sun, Hang, Jaap W. B. Bos, and Paulo Rodrigues. 2023. Destabilizing or passive? The impact of commodity index traders on equilibrium prices. *International Review of Economics and Finance* 83: 271–85. [CrossRef]

- Vigne, Samuel A., Brian M. Lucey, Fergal A. O'Connor, and Larisa Yarovaya. 2017. The financial economics of white precious metals—A survey. *International Review of Financial Analysis* 52: 292–308. [CrossRef]
- Wilson, Paul. 2015. *WPIC Platinum Quarterly Q2 2015*. London: World Platinum Investment Council.

Article

Analysis of the Impact of COVID-19 Pandemic on the Intraday Efficiency of Agricultural Futures Markets

Faheem Aslam ¹, Paulo Ferreira ^{2,3,*} and Haider Ali ¹

¹ Department of Management Sciences, Comsats University, Park Road, Islamabad 45550, Pakistan

² VALORIZA—Research Center for Endogenous Resource Valorization, 7300-555 Portalegre, Portugal

³ Instituto Politécnico de Portalegre, 7300-110 Portalegre, Portugal

* Correspondence: pferreira@ippportalegre.pt

Abstract: The investigation of the fractal nature of financial data has been growing in the literature. The purpose of this paper is to investigate the impact of the COVID-19 pandemic on the efficiency of agricultural futures markets by using multifractal detrended fluctuation analysis (MF-DFA). To better understand the relative changes in the efficiency of agriculture commodities due to the pandemic, we split the dataset into two equal periods of seven months, i.e., 1 August 2019 to 10 March 2020 and 11 March 2020 to 25 September 2020. We used the high-frequency data at 15 min intervals of cocoa, cotton, coffee, orange juice, soybean, and sugar. The findings reveal that the COVID-19 pandemic has great but varying impacts on the intraday multifractal properties of the selected agricultural future markets. In particular, the London sugar witnessed the lowest multifractality while orange juice exhibited the highest multifractality before the pandemic declaration. Cocoa became the most efficient while the cotton exhibited the minimum efficient pattern after the pandemic. Our findings show that the highest improvement is found in the market efficiency of orange juice. Furthermore, the behavior of these agriculture commodities shifted from a persistent to an antipersistent behavior after the pandemic. The information given by the detection of multifractality can be used to support investment and policy-making decisions.

Keywords: COVID-19; pandemic; agriculture; commodity; MF-DFA; high frequency; efficiency



Citation: Aslam, Faheem, Paulo Ferreira, and Haider Ali. 2022.

Analysis of the Impact of COVID-19 Pandemic on the Intraday Efficiency of Agricultural Futures Markets.

Journal of Risk and Financial Management 15: 607. <https://doi.org/10.3390/jrfm15120607>

Academic Editor: Kentaro Iwatsubo

Received: 20 October 2022

Accepted: 14 December 2022

Published: 15 December 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

According to the World Health Organization (WHO), COVID-19 has led to more than 620,000,000 confirmed infections along with more than 6,500,000 confirmed deaths. Keeping in view the alarming levels of spread and severity, the World Health Organization (WHO), on 11 March 2020, declared COVID-19 a pandemic. In addition to human loss, the indirect effects through fear and uncertainty have fostered a sense of emergency and tendency of pessimism (Barrafrem et al. 2020).

The COVID-19 pandemic has far-reaching economic implications and is becoming an extremely serious economic event (Laing 2020). It triggered a global lockdown and significantly impacted global mobility. Sadowski et al. (2021) investigated the connection between the transmission of COVID-19 and human movement and discovered that “retail and recreation areas”, “transit stations”, “workplaces”, “grocery stores”, and “pharmacies” are the hotspots for COVID-19 dissemination. Hence, the transport and energy industries suffered significant financial losses as a result of restricted movement. All forms of transport were suspended, with the exception of a few emergency situations. A similar situation prevailed in all other sectors, leading to a sharp drop in energy demand. The food business attempted to cope with COVID-19 but supply lines from farmers to retail stores and consumers were disrupted by the closure of many restaurants. However, grocery businesses and supermarkets benefited reasonably well in terms of volume of sales and earnings as a result of panic buying. Due to delays in electronic goods and industry shipments, many

technological companies ceased operations. However, COVID-19 raised the need for medical supplies and medication, which increased sales for pharmaceutical and healthcare organizations. Likewise, the revenue of Internet-based companies soared due to increased online activities such as remote office working and remote learning (Alam et al. 2021).

Global financial markets, on the other hand, also reacted strongly to this immense outbreak (Nicola et al. 2020). Fast-growing literature is focusing on different aspects of financial impacts caused by the COVID-19 pandemic. During the COVID-19 pandemic, financial markets became risky (Ali et al. 2020; Barro et al. 2020), with a drastic decline in the stock market indices (Czech et al. 2020; McKibbin and Vines 2020) and causing enormous losses (Zhang et al. 2020). This pandemic had larger impacts on the US stock market than other regions (Garcin et al. 2020). In comparison, Topcu and Gulal (2020) reported a larger impact on Asian stock markets than on European ones. From a global perspective, Aslam et al. (2020d) documented a structural change and significant changes in the financial networks due to COVID-19. Furthermore, the market's reaction to confirmed cases is larger than the confirmed deaths (Ashraf 2020). Likewise, Ali et al. (2020) reported the deterioration of financial markets when this outbreak changed from an epidemic to a pandemic.

Since the beginning of the pandemic, we can attest to a growing number of research studies focusing on the financial impacts of COVID-19 pandemic. Recently, different topics had been developed, including financial networks, (Aslam et al. 2020d; Zhang et al. 2020), stock market reactions (Aslam et al. 2020c; Zhang et al. 2020), exchange-rate fluctuation during pandemic (Njindan Iyke 2020), oil-market reactions (Apergis and Apergis 2020; Devpura and Narayan 2020), air-quality performance and multifractality (Ming et al. 2020; Sipra et al. 2021), insurance (Wang et al. 2020b), and gold and cryptocurrencies (Corbet et al. 2020).

Multifractality is central to the science of complexity and it is possible to find different applications in several essential areas of scientific activity, including physics (Muzy et al. 2008; Subramaniam et al. 2008), chemistry (Stanley and Meakin 1988; Udovichenko and Strizhak 2002), biology (Makowiec et al. 2009; Rosas et al. 2002), hydrology (Telesca et al. 2005b), environment (Farjah 2019; Sipra et al. 2021), linguistics (Drożdż et al. 2016), physiology (Nagy et al. 2017), psychology (Kelty-Stephen 2017), behavioral sciences (Ihlen and Vereijken 2013), economics (Drożdż et al. 2010), or even in music (Jafari et al. 2007) and sin markets (Aslam et al. 2021a). Particularly, multifractal detrended fluctuation analysis (MF DFA) is a stronger tool capable of detecting long-term correlations in nonstationary time series (Laib et al. 2018). In the context of finance, it helps to determine the efficiency ranking of the markets under study despite revealing the extent of the inefficiency (Rizvi et al. 2014). After the seminal works of (Mandelbrot 1967; Mandelbrot and Wallis 1969; Mandelbrot 1971, 1982, 1997), when the author introduced the concept of fractal geometry, after investigating the behavior of cotton prices and finding that these commodity prices do not exhibit a random-walk behavior, this particular method was applied to several fields with data structures, including finance (Kumar and Deo 2009; Oh et al. 2010; Podobnik and Stanley 2008; Wang and Liu 2010).

The COVID-19 pandemic also affected the efficiency of different financial markets. For instance, the intraday efficiency of European stock markets and exchange-rate markets declined during the COVID-19 outbreak (Aslam et al. 2020b). Furthermore, Aslam et al. (2020a) reported that the stock-market efficiency varies with the evolution of COVID-19 with decreasing efficiency in February–March (2020) and a recovery in April–May (2020). Particularly with respect to commodity markets, the experimental findings of (Wang et al. 2020a, 2020b) indicate that the cross-correlations of multifractality between crude oil and sugar future markets remained stronger during the pandemic. On the other hand, the efficiency of the cryptocurrency market improved during the COVID-19 pandemic (Aslam et al. 2021b; Mnif et al. 2020).

Historically, a diverse range of techniques had been adopted to analyze the efficiency of agricultural markets, including the cointegration test (Ali and Gupta 2011), VR test

(Mishra 2019), and social loss index (Consuegra and Garcia-Verdugo 2017), among others. However, few studies have investigated the efficiency of agriculture commodities by applying econophysics techniques such as DFA (Cao and Xu 2016; Kim et al. 2011; Kim and Venkatachalam 2011) and MF-DFA (Aslam et al. 2022a). Li and Lu (2011) applied MF-DFA and multifractal spectrum analysis on Chinese agricultural markets and confirmed the presence of multifractality, with highest levels in the hard winter wheat commodity market. In a recent study, Stosic et al. (2020) examined the fractal behavior of the Brazilian agriculture market through MF-DFA and confirmed that coffee futures showed the highest multifractality among the other commodities.

This paper is unique and different in many ways. Unlike other financial markets, the agricultural commodity future markets were jolted during COVID-19 through three main channels: demand, supply, and heightened uncertainty (Sifat et al. 2021). Due to basic exposure and market emotions, different commodity classes experienced varying degrees of shocks. For instance, as global demand declined during the pandemic, oil prices fell to their lowest levels since 1995, falling more steeply than other commodities and financial markets (Aslam et al. 2022b). The futures of agricultural commodities such as soybean and rice rose sharply, while corn and coffee remained stable, presumably as a result of steady real demand, declining currencies, and decreasing production of edible oil. The futures markets were established to allow economic agents to protect themselves against price risk. However, these economic and financial shocks brought about by the pandemic outbreak inflate a long-standing futures-related conundrum. Hence, it is essential to pay close attention to how commodity markets futures behaved during the COVID-19 outbreak.

We make the following contributions to the recent literature on COVID-19 and its impact on agricultural commodity markets. Firstly, as far as we are aware, this is the first study to distinguish between the impacts of the COVID-19 pandemic on the efficiency of agricultural futures markets in a multifractal context. To do this, we followed (Aslam et al. 2021b), and divided the overall data into two equal periods, before the outbreak of the pandemic (1 August 2018 to 10 March 2020) and afterwards (11 March 2020 to 25 September 2020). Secondly, unlike other studies, we employed the 15 min high-frequency data of six agricultural futures: cocoa, cotton, coffee, orange juice, soybean, and sugar. This is justified by the fact that, compared to daily data, high-frequency data allows for a more precise calculation of the complexity and multifractal properties in a time series (Aslam et al. 2021b, 2021c). Thirdly, the robust fractal market-theory-based MF DFA of Kantelhardt et al. (2002) is used in this study. Using this approach, we investigate four key questions: (1) Do agricultural future markets have multifractal properties? (2) Do agricultural future markets' multifractal characteristics change before and after the declaration of the COVID-19 pandemic? (3) Does the strength of the multifractal characteristics change before and after the declaration of the COVID-19 pandemic? (4) Do the persistent behavior or autocorrelation characteristics change before and after the declaration of the COVID-19 pandemic?

The rest of the paper proceeds as follows. Data and methodology are discussed in Section 2. Section 3 presents the methodology used in this paper, while results are presented in Section 4. The final section summarizes the key conclusions and recommendations of the paper.

2. Data Description

Despite the fear of inciting panic or prompting some countries to flag in their efforts, the World Health Organization (WHO) had to declare the novel coronavirus (COVID-19) as a global pandemic on 11 March 2020 (Maier and Brockmann 2020). In a few weeks, this pandemic shaved off nearly a third of the global market capitalization. Numerous studies have revealed that COVID-19 primarily impacted financial markets in the second and third quarters of 2020, which is from March 2020 to September 2020 (Aslam et al. 2021b). For instance, Aslam et al. (2021b) quantified the self-similarity intensity of six stock markets in Europe and Asia and investigated the quarterly variations in herding behavior using MF DFA. They discovered that COVID-19 had a variable influence on these markets on a

quarterly basis, but that the impact peaked in the second quarter and recovered in the third and fourth quarters of 2020. Hence, following Aslam et al. (2021b), we restrict our period from 1 August 2019 to 25 September 2020 and split the dataset into two equal periods of seven months. By using the pandemic declaration date, the intraday prices ranging from 1 August 2019 to 10 March 2020 refer to the period before the pandemic and prices from 11 March 2020 to 25 September 2020 form the sample for the period after announcing the pandemic. To overcome the problem of microstructure noise and duplicate values, the high-frequency data is aggregated into a 15 min frequency, which is the highest and best frequency and is in line with (Aslam et al. 2021b; Chen et al. 2022; Zhang and Ma 2021). Due to removal of the duplicate values and different number of trading days in different months, we have a different number of observations for these time periods. The list of agriculture commodities, data range, and the number of observations of both time periods are presented in Table 1. These commodities were chosen solely on the basis of the availability of intraday data. In Figure 1, the trend shows an instant decline in all commodity prices after the declaration of the pandemic, except for orange Juice.

Table 1. Data description of agriculture commodities (1 August 2019 to 25 September 2020).

S. No.	Commodity	Number Obs. Before Pandemic (1 August 2019 to 10 March 2020)	Number Obs. After Pandemic (11 March 2020 to 25 September 2020)
1	US Cocoa	8986	4704
2	US Coffee	5385	4999
3	US Cotton	10,143	9567
4	US Orange Juice	3460	3314
5	US Soybean	9414	8961
6	London Sugar	5421	5014

One of the reasons for this was an increase in the demand of orange juice across the globe. For the return rates of the agriculture commodity futures, we used the normal logarithm difference, i.e.,

$$r(t) = \ln p(t) - \ln p(t - 1) \tag{1}$$

where $p(t)$ is the price of a given index at time t . It is clear in Figure 2 that after the pandemic declaration on 11 March 2020, more volatility returns can be observed in all commodities.

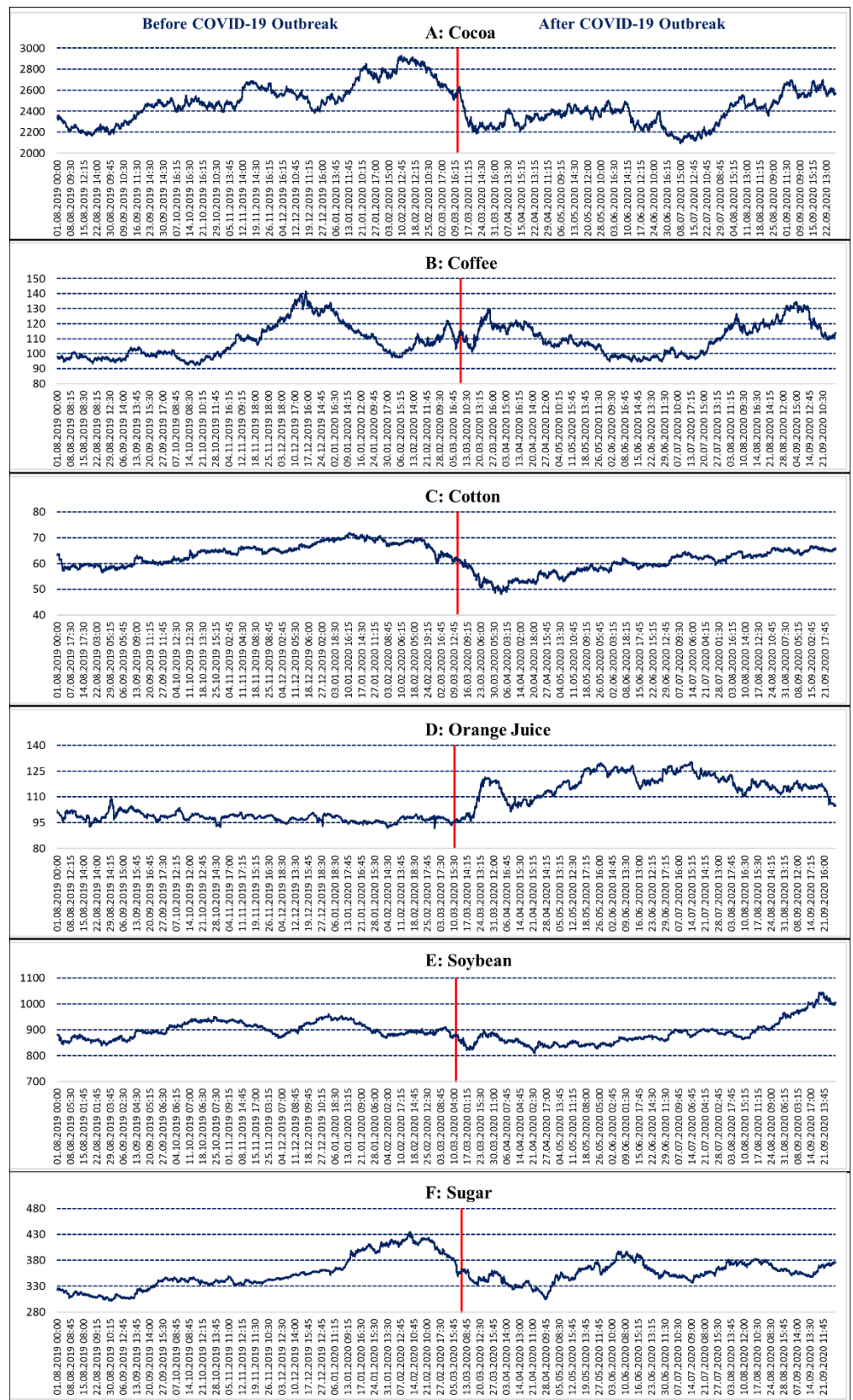


Figure 1. Time trend of 15 min prices from 1 August 2019 to 25 September 2020.

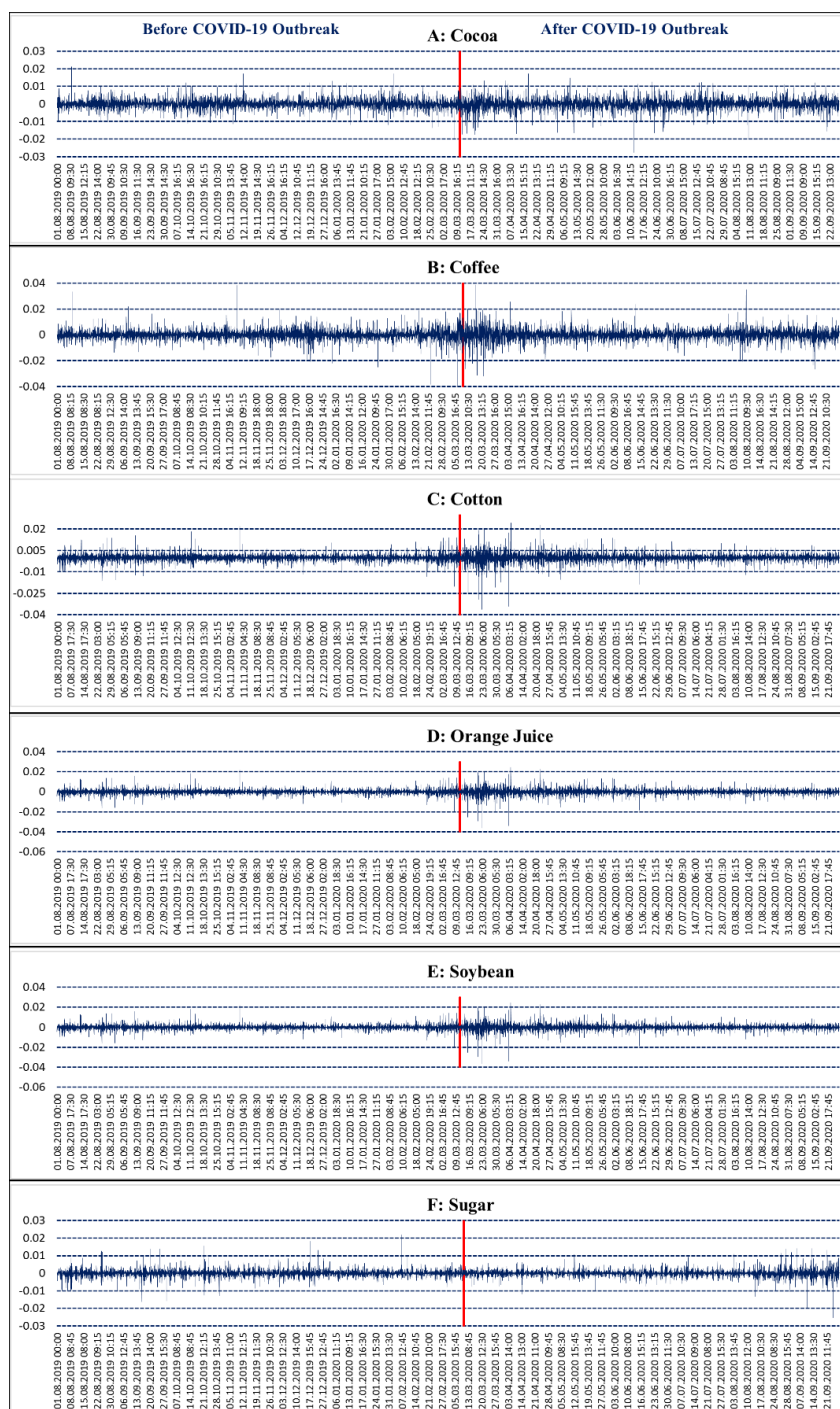


Figure 2. Time trend of 15 min log returns from 1 August 2019 to 25 September 2020.

3. Multifractal Detrended Fluctuation Analysis (MF-DFA)

The procedure of MF-DFA detailed by Kantelhardt et al. (2002) could be briefly identified as follows.

Considering a time series $\{Z_t, t=1, \dots, N\}$ with length N , the first step considers the construction of the profile $X(k)$, i.e.,

$$X(k) = \sum_{t=1}^k (z_t - \bar{z}) \tag{2}$$

where $k = 1, 2, 3 \dots, N$ and \bar{z} is the observed mean. This profile obtained is divided in $N_s = \text{int} \frac{N}{s}$ nonoverlapping boxes with a length s . According to Kantelhardt et al. (2002), it is possible to use a final part of the sample, when N is a nonmultiple of s , allowing to have a total of $2N_s$ segments, and following the rule proposed by Peng et al. (1994), i.e., $10 < s < \frac{N_s}{5}$. After this, a local trend is estimated for all $2N_s$ segments based on the ordinary least squares, with

$$F^2(s, m) = \frac{1}{s} \sum_{j=1}^s \{X[(m-1)s + j] - x_m(j)\}^2 \tag{3}$$

for the segments between $m = 1$ and $m = N_s$, while for the segments between $m = N_{s+1}$ and $m = 2N_s$, this is given by

$$F^2(s, m) = \frac{1}{s} \sum_{j=1}^s \{X[N - (m - N_s)s + j] - x_m(j)\}^2 \tag{4}$$

representing $x_m(j)$ as the polynomial fit of the segments. Prior functions are averaged, allowing to obtain the q th-order fluctuation function for any $q \neq 0$, given by

$$F_{q(s)} = \left\{ \frac{1}{2N_s} \sum_{\rho=1}^{2N_s} [F^2(s, \rho)]^{q/2} \right\}^{1/q} \tag{5}$$

and for $q = 0$

$$F_{0(s)} = \exp \left\{ \frac{1}{4N_s} \sum_{\rho=1}^{2N_s} \ln [F^2(s, \rho)] \right\} \tag{6}$$

Negative values of q represent small fluctuations and positive values represent larger fluctuations. The value at $q = 2$ represents the DFA exponent. The value of $F_{q(s)}$ increases with s , and performing a log–log regression allows to obtain a power law given by

$$F_{q(s)} \sim S^{h(q)} \tag{7}$$

The value of $h(q)$ is the Hurst exponent, which measures the dependence levels of financial assets (see, for example, (Domino 2011; Pleşoianu et al. 2012)). A value of $h(q) = 0.5$ implies that the financial market under analysis behaves like a random walk; if $h(q) > 0.5$ and $h(q) < 0.5$, that market represents, respectively, persistent and antipersistent patterns.

The Renyi exponent $\tau(q)$ can be calculated by using the value of $h(q)$ to measure the mono- or multifractality behavior of a given series.

$$\tau(q) = qh(q) - 1 \tag{8}$$

By using the Legendre transformation, the relationship between $\tau(q)$ and $f(\alpha)$ can be written as follows:

$$f(\alpha) = q\alpha - \tau(q) \tag{9}$$

Through Equations (8) and (9), we can obtain the multifractal spectrum $f(\alpha)$

$$\alpha = h(q) - qh'(q) \tag{10}$$

$$f(\alpha) = q(\alpha - h(q)) + 1 \tag{11}$$

with α being the Holder exponent given by

$$\alpha = h_q + q \frac{\gamma h_q}{\gamma_q} - \tau_q \tag{12}$$

The degree of multifractality can be quantified by the width of spectrum $\Delta\alpha$ and Δh , i.e.,

$$\Delta h = h_{(qmin)} - h_{(qmax)} \tag{13}$$

or

$$\Delta\alpha = \alpha_{(max)} - \alpha_{(min)} = h(-\infty) - h(+\infty) = \frac{\ln(\frac{a}{b})}{\ln 2} \tag{14}$$

A higher width of $\Delta\alpha$ or Δh implies a higher degree of multifractality.

4. Empirical Findings

The estimation of the MF-DFA for each commodity at each period under analysis provides us the information given in Figure 3, in this case representing the estimations for the cocoa. The panel (A) (top left) shows the log–log relationship between the fluctuation function (F_q) and the time scale s from $q = -30$ to $q = 30$, which, in the case under analysis, exhibits a well-fit straight line. In panel (B), we have the generalized Hurst exponent $h(q)$, which is a decreasing function of the value of q . Panel (A) has the quickest decline, which will be consistent with the difference of the maximum and minimum values of $h(q)$. On the other hand, Δh indicates the higher multifractality levels and is related with higher degrees of inefficiency. In panel (A), we have the Renyi exponent, which is nonlinear in the case of the existence of multifractality, and panel (D) represents the multifractal spectrum, which identifies multifractality when it is described by a single-humped shape. In conclusion, Figure 3 shows that there exists evidence of multifractality in cocoa both before and after the declaration of the pandemic.

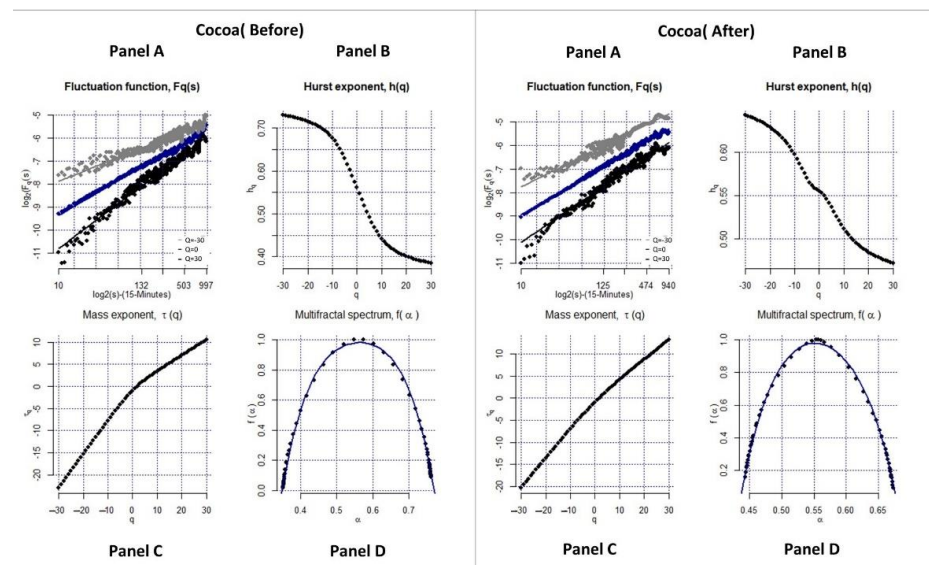


Figure 3. MF-DFA results for cocoa. (Panel A) Fluctuation functions for $q = [-30,0,+30]$; (Panel B) Generalized Hurst exponent depending on q ; (Panel C) Mass exponent $t(q)$; (Panel D) Multifractal spectrum. On the left is the information for the period before the declaration of the COVID-19 pandemic and on the right is the information for the period after the declaration of the COVID-19 pandemic.

For the behavior of the remaining commodities, the evidence is rather similar, excepting the case of the orange juice in the period after the pandemic (see Figures 4 and 5, showing all the other results). In that case, all the panels show differences when compared with the other general figures. In panel (A), the log–log relationship shows some instability; in panel (B), the Hurst exponent decreases sharply from $q = -30$ to $q = 30$, then goes up until $q = 5$, and decreases again until $q = 30$; panel (C) is near to a linear relationship, but

panel (D), representing the multifractal spectrum $f(\alpha)$, also has a single humped-shape, indicating that the series is multifractal.

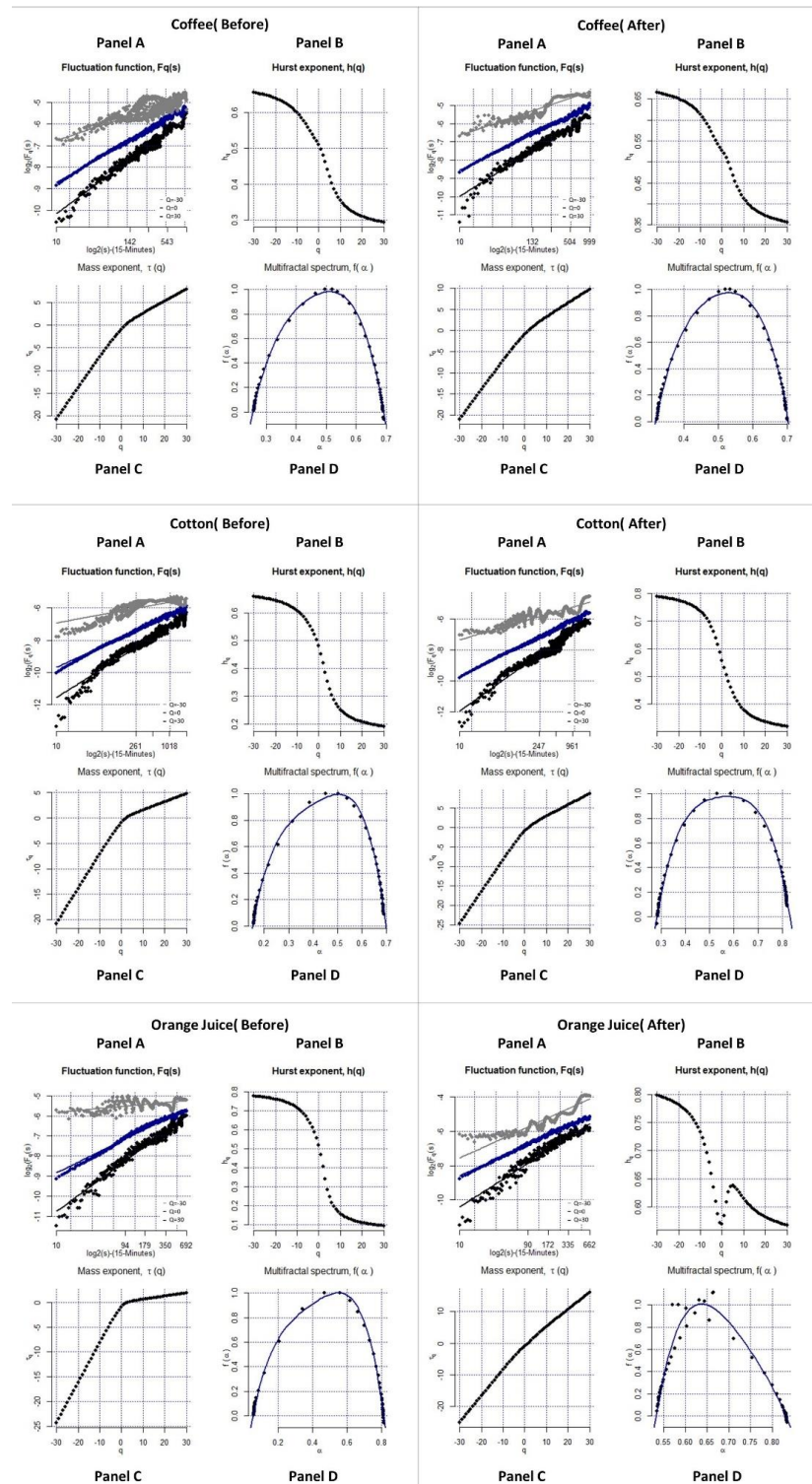


Figure 4. MFDA results for coffee, cotton, and orange Juice. (Panel A) Fluctuation functions for $q = [-30,0,+30]$; (Panel B) Generalized Hurst exponent depending on q ; (Panel C) Mass exponent $t(q)$; (Panel D) Multifractal spectrum. On the left is the information for the period before the declaration of the COVID-19 pandemic and on the right is the information for the period after the declaration of the COVID-19 pandemic.

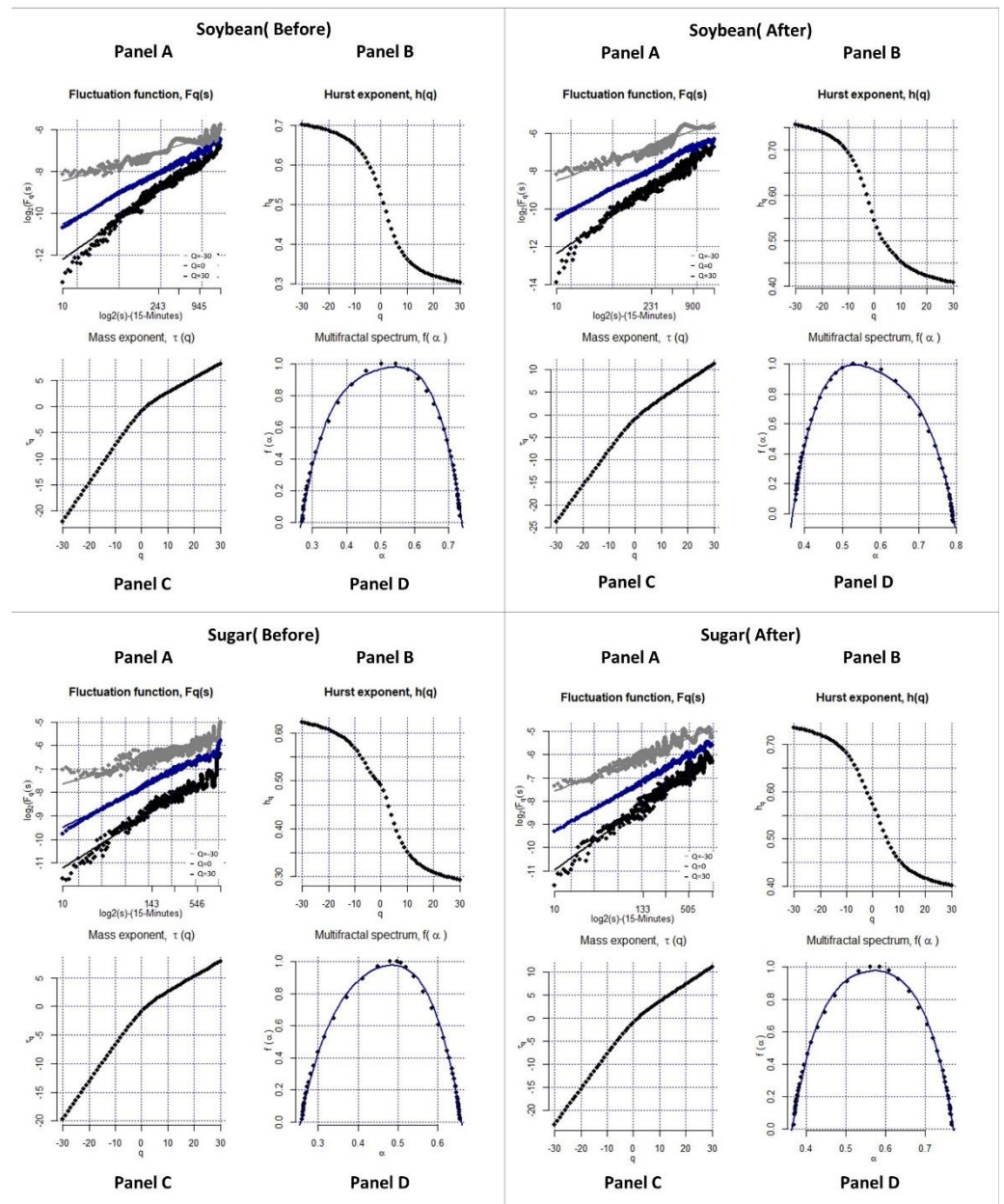


Figure 5. MFDDFA results for soybean and sugar. (Panel A) Fluctuation functions for $q = [-30,0,+30]$; (Panel B) Generalized Hurst exponent depending on q ; (Panel C) Mass exponent $t(q)$; (Panel D) Multifractal spectrum. On the left is the information for the period before the declaration of the COVID-19 pandemic and on the right is the information for the period after the declaration of the COVID-19 pandemic.

The generalized Hurst exponents of all commodities future markets over the range of $q \in [-30, 30]$ are shown in Table 2. All the $h(q)$ values for ‘before pandemic’ and ‘after pandemic’ decrease with the increase of q , confirming that the generalized Hurst exponent $h(q)$ depends on the value of q , suggesting the presence of multifractality in these markets for both periods.

Table 2. Generalized Hurst exponents for the agriculture futures ranging from $q = -30$ to $q = 30$.

Q	Before COVID-19 Outbreak						After COVID-19 Outbreak					
	Cocoa	Coffee	Cotton	Orange Juice	Soybean	Sugar	Cocoa	Coffee	Cotton	Orange Juice	Soybean	Sugar
−30	0.7302	0.6568	0.6588	0.7765	0.7007	0.6219	0.6424	0.6658	0.7878	0.7983	0.7559	0.7346
−29	0.7292	0.6556	0.6578	0.7754	0.6996	0.6208	0.6414	0.6648	0.7868	0.7971	0.7547	0.7336
−28	0.7281	0.6544	0.6567	0.7741	0.6985	0.6196	0.6403	0.6636	0.7857	0.7959	0.7535	0.7324
−27	0.7269	0.653	0.6555	0.7728	0.6973	0.6184	0.6391	0.6624	0.7846	0.7945	0.7522	0.7312
−26	0.7256	0.6516	0.6543	0.7714	0.696	0.617	0.6379	0.661	0.7833	0.7931	0.7507	0.7299
−25	0.7243	0.6501	0.6529	0.7698	0.6946	0.6156	0.6366	0.6596	0.782	0.7915	0.7492	0.7285
−24	0.7228	0.6485	0.6515	0.7682	0.6931	0.6141	0.6352	0.6581	0.7805	0.7898	0.7475	0.7269
−23	0.7212	0.6467	0.6499	0.7664	0.6915	0.6124	0.6337	0.6565	0.779	0.788	0.7457	0.7253
−22	0.7195	0.6448	0.6482	0.7644	0.6898	0.6106	0.6321	0.6547	0.7773	0.786	0.7437	0.7235
−21	0.7176	0.6427	0.6464	0.7623	0.688	0.6086	0.6303	0.6528	0.7754	0.7838	0.7416	0.7216
−20	0.7155	0.6404	0.6444	0.76	0.6859	0.6065	0.6284	0.6507	0.7734	0.7814	0.7392	0.7195
−19	0.7133	0.6379	0.6422	0.7574	0.6837	0.6042	0.6264	0.6484	0.7713	0.7787	0.7366	0.7172
−18	0.7108	0.6352	0.6399	0.7546	0.6813	0.6016	0.6241	0.6459	0.7688	0.7758	0.7338	0.7147
−17	0.7081	0.6323	0.6373	0.7515	0.6786	0.5988	0.6217	0.6432	0.7662	0.7725	0.7307	0.7119
−16	0.7051	0.629	0.6344	0.748	0.6757	0.5957	0.6191	0.6401	0.7632	0.7688	0.7272	0.7089
−15	0.7017	0.6254	0.6312	0.7441	0.6724	0.5923	0.6162	0.6368	0.7599	0.7646	0.7233	0.7055
−14	0.6979	0.6214	0.6277	0.7398	0.6688	0.5886	0.613	0.6331	0.7562	0.7599	0.719	0.7017
−13	0.6937	0.6169	0.6237	0.7348	0.6648	0.5843	0.6095	0.629	0.752	0.7545	0.7141	0.6975
−12	0.6889	0.612	0.6193	0.7292	0.6603	0.5796	0.6057	0.6244	0.7472	0.7483	0.7085	0.6927
−11	0.6834	0.6064	0.6142	0.7227	0.6552	0.5743	0.6015	0.6193	0.7417	0.7411	0.7022	0.6873
−10	0.6771	0.6002	0.6085	0.7152	0.6494	0.5682	0.597	0.6135	0.7352	0.7327	0.6949	0.6811
−9	0.6699	0.5933	0.602	0.7066	0.6429	0.5615	0.592	0.607	0.7277	0.7226	0.6865	0.674
−8	0.6617	0.5857	0.5946	0.6964	0.6353	0.5538	0.5867	0.5996	0.7188	0.7107	0.6768	0.6659
−7	0.6521	0.5773	0.586	0.6844	0.6267	0.5453	0.581	0.5914	0.7081	0.6964	0.6655	0.6566
−6	0.6412	0.5682	0.576	0.6702	0.6169	0.5359	0.5753	0.5823	0.6953	0.6792	0.6524	0.6461
−5	0.629	0.5587	0.5646	0.6536	0.6056	0.5262	0.5697	0.5724	0.6797	0.6587	0.6373	0.6343
−4	0.6158	0.549	0.5517	0.6343	0.5929	0.5169	0.5649	0.562	0.661	0.635	0.6203	0.6217
−3	0.6021	0.5393	0.5373	0.6122	0.5785	0.509	0.5611	0.5517	0.6389	0.6098	0.6018	0.609
−2	0.5884	0.5298	0.5215	0.587	0.5624	0.5032	0.5586	0.5425	0.6137	0.5868	0.5826	0.5968
−1	0.5748	0.5203	0.5032	0.5572	0.5443	0.4982	0.5569	0.5346	0.5862	0.5713	0.5633	0.5852
0	0.5612	0.51	0.4807	0.5195	0.5242	0.491	0.5553	0.5275	0.5577	0.5694	0.5446	0.5736
1	0.5474	0.4976	0.4525	0.4695	0.5027	0.4796	0.5531	0.5203	0.5302	0.5842	0.528	0.5611
2	0.5332	0.4816	0.4194	0.4059	0.4804	0.4645	0.55	0.5111	0.5049	0.6079	0.5144	0.5472
3	0.519	0.462	0.3846	0.3398	0.4585	0.4468	0.5461	0.4988	0.4816	0.6268	0.5039	0.5325
4	0.5051	0.4408	0.3525	0.2855	0.438	0.4283	0.5414	0.484	0.4603	0.636	0.4952	0.5177
5	0.4917	0.4203	0.3255	0.2458	0.4198	0.4106	0.5361	0.4686	0.4412	0.6377	0.4874	0.5038
6	0.4793	0.4022	0.3036	0.2173	0.404	0.3949	0.5307	0.4543	0.4246	0.6352	0.4799	0.4913
7	0.4681	0.3867	0.2862	0.1963	0.3907	0.3814	0.5255	0.4417	0.4106	0.6306	0.4728	0.4802
8	0.4582	0.3738	0.2722	0.1804	0.3794	0.3698	0.5205	0.4308	0.3987	0.6253	0.4661	0.4706
9	0.4495	0.363	0.2609	0.1679	0.3698	0.3601	0.5158	0.4215	0.3886	0.6199	0.46	0.4622
10	0.4418	0.354	0.2515	0.1578	0.3616	0.3517	0.5115	0.4135	0.38	0.6147	0.4545	0.455
11	0.4351	0.3463	0.2437	0.1495	0.3546	0.3446	0.5076	0.4066	0.3726	0.6099	0.4495	0.4488
12	0.4293	0.3397	0.2371	0.1426	0.3486	0.3384	0.5041	0.4006	0.3662	0.6055	0.445	0.4433
13	0.4241	0.3341	0.2314	0.1367	0.3433	0.3331	0.5008	0.3954	0.3607	0.6015	0.441	0.4385
14	0.4196	0.3291	0.2265	0.1316	0.3386	0.3284	0.4979	0.3909	0.3558	0.5978	0.4374	0.4343
15	0.4155	0.3248	0.2222	0.1271	0.3345	0.3242	0.4951	0.3868	0.3515	0.5945	0.4341	0.4305
16	0.4119	0.3209	0.2184	0.1232	0.3309	0.3205	0.4926	0.3832	0.3477	0.5914	0.4312	0.4272
17	0.4086	0.3175	0.215	0.1197	0.3276	0.3172	0.4903	0.38	0.3443	0.5887	0.4285	0.4242
18	0.4057	0.3145	0.212	0.1166	0.3246	0.3143	0.4882	0.3771	0.3412	0.5862	0.4261	0.4215
19	0.4031	0.3117	0.2093	0.1138	0.3219	0.3116	0.4863	0.3745	0.3384	0.5839	0.4238	0.419
20	0.4006	0.3092	0.2069	0.1113	0.3195	0.3092	0.4845	0.3721	0.3359	0.5817	0.4218	0.4168
21	0.3984	0.3069	0.2047	0.109	0.3173	0.307	0.4828	0.3699	0.3336	0.5798	0.4199	0.4147
22	0.3964	0.3049	0.2026	0.1069	0.3152	0.305	0.4812	0.3679	0.3315	0.578	0.4181	0.4129
23	0.3945	0.303	0.2008	0.105	0.3134	0.3031	0.4798	0.3661	0.3296	0.5763	0.4165	0.4111
24	0.3928	0.3012	0.1991	0.1032	0.3116	0.3014	0.4784	0.3644	0.3278	0.5748	0.415	0.4096
25	0.3912	0.2996	0.1975	0.1016	0.3101	0.2998	0.4771	0.3629	0.3262	0.5733	0.4137	0.4081

Table 2. Cont.

Q	Before COVID-19 Outbreak						After COVID-19 Outbreak					
	Cocoa	Coffee	Cotton	Orange Juice	Soybean	Sugar	Cocoa	Coffee	Cotton	Orange Juice	Soybean	Sugar
26	0.3897	0.2981	0.196	0.1001	0.3086	0.2984	0.476	0.3615	0.3247	0.572	0.4124	0.4067
27	0.3884	0.2967	0.1947	0.0987	0.3072	0.297	0.4748	0.3601	0.3233	0.5707	0.4112	0.4054
28	0.3871	0.2954	0.1934	0.0974	0.3059	0.2957	0.4738	0.3589	0.322	0.5696	0.41	0.4043
29	0.3859	0.2942	0.1922	0.0962	0.3047	0.2946	0.4728	0.3577	0.3207	0.5685	0.409	0.4031
30	0.3848	0.2931	0.1911	0.0951	0.3036	0.2935	0.4719	0.3566	0.3196	0.5674	0.408	0.4021
Delta H	0.3454	0.3637	0.4677	0.6814	0.3971	0.3284	0.1705	0.3092	0.4682	0.2309	0.3479	0.3325
Delta α	0.4087	0.4324	0.531	0.75	0.4624	0.3942	0.2291	0.3751	0.5341	0.2982	0.4136	0.3977

For instance, before the pandemic, the highest value of $h(q)$ of London sugar, represented in Table 2, is 0.62 for $q = -30$, falling to 0.29 at $q = 0$, and with the lowest value of 0.38 for $q = 30$. Likewise, after the pandemic declaration, the highest value of $h(q)$ of USA cocoa is 0.64 for $q = -30$, falling to 0.55 at $q = 0$, and with the lowest value of 0.47 for $q = 30$. Comparable declining patterns are found in other markets for both time periods.

The width of the generalized Hurst exponents Δh is documented in Table 2. The value of Δh specifies the degree of the multifractality, with a smaller width meaning lower multifractality levels (Telesca et al. 2005a).

By comparing $\Delta h(q)$ in Table 2 in subperiod 1 of before the pandemic, the orange juice future market shows the highest multifractality patterns, with $\Delta h = 0.68$, followed by cotton ($\Delta h = 0.47$), while the sugar ($\Delta h = 0.33$) and cocoa ($\Delta h = 0.34$) markets show the lowest multifractality, respectively. After the pandemic declaration on 11 March 2020, a significant change in the multifractality can be observed. After the pandemic declaration, the cotton future market exhibited the highest multifractality ($\Delta h = 0.47$), followed by soybeans ($\Delta h = 0.35$), while the cocoa market showed the lowest degree of multifractality with multifractality ($\Delta h = 0.17$). Overall, an increase in the multifractally Δh is confirmed in two out of six (sugar and cotton) markets, while a decline is found in the remaining four markets. The comparative multiple spectra of all six agriculture future markets are presented in Figure 6.

The same conclusion can also be obtained by comparing the spectrum width $\Delta \alpha$ in Figure 6 and last row of Table 2. Finally, the COVID-19 pandemic significantly affected the persistence level of commodity future markets. Before the pandemic, only the USA cocoa future market ($h(q) = 0.53$) showed persistent behavior, while the remaining five markets showed an antipersistence feature. However, all the markets ($h(q) > 0.50$) exhibit evidently persistent features with a clear shape and a trace of herd behavior. Somewhat similar patterns are found in the bitcoin market during the COVID-19 pandemic (Aslam et al. 2021b; Mnif et al. 2020).

The multifractal characteristics reveal the efficiency of financial markets (Anagnostidis et al. 2016). Consequently, based on the multifractal properties examined, USA cotton and London sugar showed a decline level while cocoa, coffee, orange juice, and soybean showed an efficiency improvement after the declaration of the COVID-19 pandemic. The market efficiency of orange juice improved the most among these six commodities, as it was the most inefficient market before the pandemic and became relatively more efficient after the pandemic. Overall, the findings reveal that the COVID-19 pandemic has a great but varying impact on the multifractal properties and persistence levels of agricultural future markets.

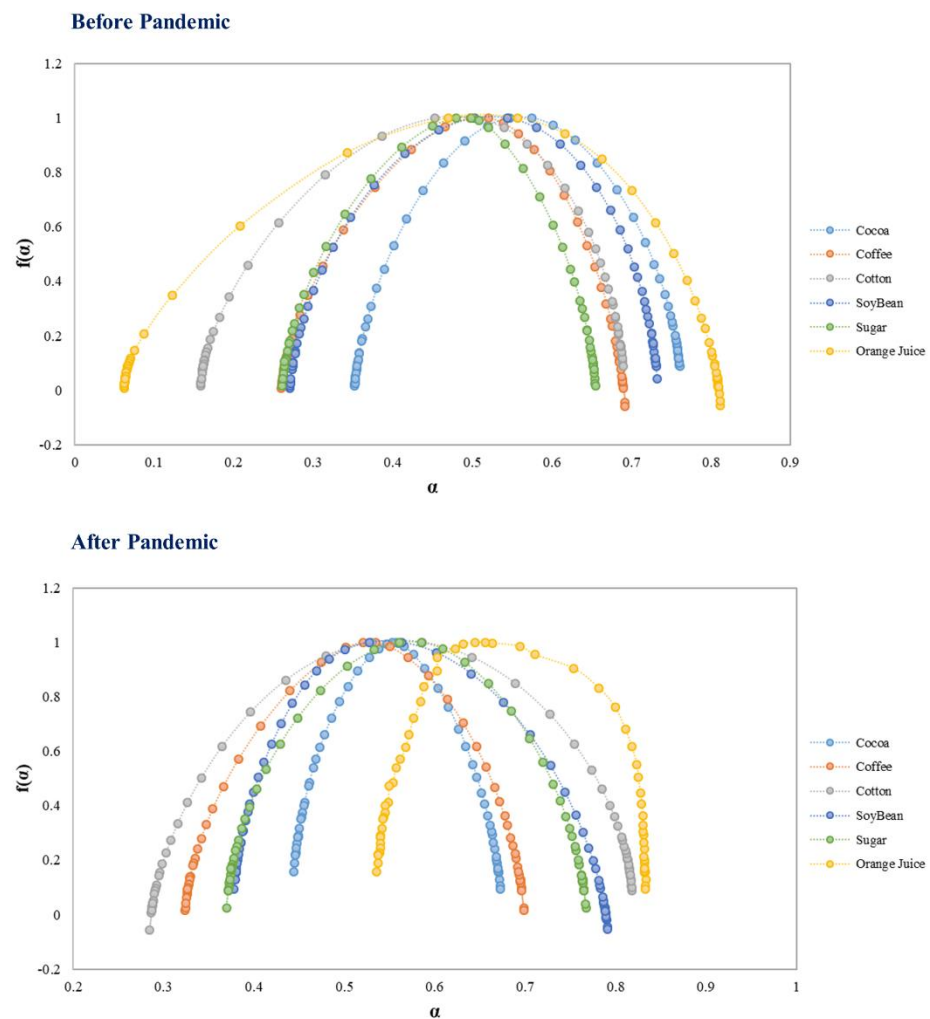


Figure 6. The multifractal spectra for agricultural commodities before and after the declaration of the pandemic.

5. Conclusions and Discussion

Agricultural commodities are viewed as central resources for the food security and social stability of any economy. With this background, this paper investigates why it is critical to monitor their prices, especially during times of increased instability. Additionally, during times of global uncertainty, commodities are increasingly seen as replacement assets by investors seeking alternative investments for diversification purposes. Goodell (2020) describes this pandemic as a worldwide crisis with severe economic harm that has never been experienced before. This period of economic and social unrest, which is akin to the GFC of 2007–2008, stands out, and much research has concentrated on analyzing the effects of the COVID-19 pandemic on various commodity markets (Lu and Zeng 2022; Sifat et al. 2021; Štreimikienė et al. 2022; Wang et al. 2020a). This, however, does not make this paper any less valuable. The most significant aspect of this paper is that it splits a 14-month dataset into two equal periods from 1 August 2019 to 10 March 2020 and 11 March 2020 to 25 September 2020. The purpose of doing this is to better understand the inner dynamics of agricultural commodity markets during an extreme crisis time, i.e., March 2020 to September 2020, as found by (Aslam et al. 2021b). They discovered that this pandemic mostly affected financial markets in the second quarter of 2020, with the third and subsequent quarters being the period of recovery. In this connection, we investigated the intraday multifractal behavior of six agricultural futures: USA cocoa, USA cotton, USA coffee, USA orange juice, USA soybean, and London sugar.

The findings of this study are as follows: Firstly, all agricultural commodities are found to have significant multifractal patterns in both periods, i.e., before and after the outbreak of COVID-19. Secondly, a significant change in the multifractality is observed, where the multifractality of cotton and sugar increases significantly after the pandemic, while cocoa, orange juice, and soybean are found to be less multifractal after the pandemic. This indicates that the market efficiency of cotton and sugar decreases significantly while the market efficiency of cocoa, orange juice and soybean increase after the pandemic. Interestingly, the overall results found orange juice to be highly efficient among all commodities. Thirdly, all the markets have $h(q)$ values greater than 0.5, indicating an increase in the strength of positive persistence after the COVID-19 pandemic.

Furthermore, our findings are in line with previous studies (Heng et al. 2020; Umar et al. 2021; Wang et al. 2020a). For instance, Wang et al. (2020a) used multifractal detrended cross-correlation analysis (MF-DCCA) to investigate the effects of the COVID-19 pandemic on the cross-correlations between the crude oil and agricultural futures markets. They discovered that, with the exception of the orange juice future market, all agricultural futures' cross-correlations rose following the outbreak of the COVID-19 pandemic. Similarly, Akyildirim et al. (2022) discovered that, during pandemics, orange juice sentiments are the least correlated with other market sentiments. The possible explanation is that orange juice was seen as a popular commodity during the pandemic, and many consumers believed that orange juice would keep them healthy during the worrying times of the COVID-19 pandemic. According to Heng et al. (2020), the dollar sales of orange juice at grocery stores rose by 50.7% for the four-week period in March–April 2020 in comparison to the dollar sales of March–April 2019. Similarly, in response to consumers' altered purchasing patterns, the futures price for frozen concentrated orange juice increased by 24 percent in mid-March 2020, making orange juice one of the best-performing commodities during the COVID-19 pandemic (Heng et al. (2020).

The existence of multifractality in agricultural commodities suggests that prices have a pattern and that price volatility has clusters, demonstrating some predictability. This, however, defies the principles of the Efficient Market Hypothesis (EMH) of Fama (1970) and presents an exploitable opportunity for investors to outperform the market and attain abnormal returns. Investors can therefore take this market persistence level into account when developing their investment strategy. Numerous studies have employed multifractality and the Hurst exponent to examine the market's inefficiency to predict changes in financial market movements (Ali et al. 2021; Alvarez-Ramirez et al. 2008; Aslam et al. 2021a, 2022a). According to our findings, a wider multiple spectrum and a wider range of Δh are the signs of greater inefficiency (Cajueiro et al. 2009; Caraianni 2012). These findings also give authorities better understanding, since multifractality brought about by market complexity or the lesser development of these markets may cause these inefficiencies (Aslam et al. 2021a; Rizvi et al. 2014).

Hence, our study has different implications for investors, practitioners, managers, and policymakers. Firstly, the multifractality caused by market complexity or by the lesser development of these markets may cause these inefficiencies (Aslam et al. 2021a; Rizvi et al. 2014). This raises the likelihood of herding behavior and irrational speculation, both of which could skew the expected payoffs from underlying investment evaluation. Consequently, these findings offer useful information for investors, practitioners, and managers in obtaining abnormal returns. Secondly, these findings help authorities and policymakers to implement measures that increase transparency, which will boost the efficiency of these futures markets, particularly during moments of economic and financial instability. Thirdly, our findings imply that investors need to be more cautious when underestimating the risk exposure to agricultural commodities, contrarily to a prevalent misconception that these commodities provide safe-haven benefits in crisis moments. Finally, it is essential to use nonlinear approaches to detect and identify all the possible patterns in the financial markets.

Despite the findings, it is relevant to identify the possible limitations of this paper, namely, the fact that it only examines the multifractal properties of agricultural futures markets before and after the COVID-19 outbreak without the identification of any conclusions about the causes which underlie the physical mechanism or multifractal properties. Hence, future studies should employ various multifractal approaches to understand the underlying sources of multifractality in these markets. This study also does not take any grain-based commodities due to data availability constraints; hence, the inclusion of more agricultural futures markets should be considered. Moreover, the analysis could also be extended by examining the impact of geopolitical risk and economic-policy-uncertainty indicators on the efficiency of these markets. Finally, the application of the rolling window on the multifractality would also provide more insights regarding the efficiency of these markets over time.

Author Contributions: Conceptualization, F.A. and P.F.; methodology, F.A. and P.F.; formal analysis, F.A. and P.F.; writing—original draft preparation, F.A. and P.F.; writing—review and editing, F.A., P.F. and H.A. All authors have read and agreed to the published version of the manuscript.

Funding: Paulo Ferreira is pleased to acknowledge financial support from Fundação para a Ciência e a Tecnologia (grant UIDB/05064/2020).

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

Acknowledgments: We sincerely appreciate all valuable comments and suggestions from the anonymous reviewers and the Academic Editor, which helped us to improve the quality of the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

- Akyildirim, Erdinc, Oguzhan Cepni, Linh Pham, and Gazi Salah Uddin. 2022. How connected is the agricultural commodity market to the news-based investor sentiment? *Energy Economics* 113: 106174. [CrossRef]
- Alam, Md Mahmudul, Haitian Wei, and Abu N. M. Wahid. 2021. COVID-19 outbreak and sectoral performance of the Australian stock market: An event study analysis. *Australian Economic Papers* 60: 482–95. [CrossRef]
- Ali, Haider, Faheem Aslam, and Paulo Ferreira. 2021. Modeling Dynamic Multifractal Efficiency of US Electricity Market. *Energies* 14: 6145. [CrossRef]
- Ali, Jabir, and Kriti Bardhan Gupta. 2011. Efficiency in agricultural commodity futures markets in India: Evidence from cointegration and causality tests. *Agricultural Finance Review* 71: 162–78. [CrossRef]
- Ali, Mohsin, Nafis Alam, and Syed Aun R. Rizvi. 2020. Coronavirus (COVID-19)—An epidemic or pandemic for financial markets. *Journal of Behavioral and Experimental Finance* 27: 100341. [CrossRef]
- Alvarez-Ramirez, Jose, Jesus Alvarez, and Eduardo Rodriguez. 2008. Short-term predictability of crude oil markets: A detrended fluctuation analysis approach. *Energy Economics* 30: 2645–56. [CrossRef]
- Anagnostidis, Panagiotis, Christos Varsakelis, and Christos J Emmanouilides. 2016. Has the 2008 financial crisis affected stock market efficiency? The case of Eurozone. *Physica A: Statistical Mechanics and its Applications* 447: 116–28. [CrossRef]
- Apergis, Emmanuel, and Nicholas Apergis. 2020. Can the COVID-19 pandemic and oil prices drive the US Partisan Conflict Index. *Energy Research Letters* 1: 13144. [CrossRef]
- Ashraf, Badar Nadeem. 2020. Stock markets' reaction to COVID-19: Cases or fatalities? *Research in International Business and Finance* 54: 101249. [CrossRef]
- Aslam, Faheem, Francisca Nogueiro, Mariana Brasil, Paulo Ferreira, Khurram Shahzad Mughal, Beenish Bashir, and Saima Latif. 2020a. The footprints of COVID-19 on Central Eastern European stock markets: An intraday analysis. *Post-Communist Economies* 33: 751–69. [CrossRef]
- Aslam, Faheem, Ijaz Ali, Fahd Amjad, Haider Ali, and Inza Irfan. 2022a. On the inner dynamics between Fossil fuels and the carbon market: A combination of seasonal-trend decomposition and multifractal cross-correlation analysis. *Environmental Science and Pollution Research*, 1–19. [CrossRef] [PubMed]
- Aslam, Faheem, Paulo Ferreira, Fahd Amjad, and Haider Ali. 2021a. The Efficiency of Sin Stocks: A Multifractal Analysis of Drug Indices. *The Singapore Economic Review* 1–22. [CrossRef]
- Aslam, Faheem, Paulo Ferreira, Haider Ali, and Ana Ercília José. 2022b. Application of Multifractal Analysis in Estimating the Reaction of Energy Markets to Geopolitical Acts and Threats. *Sustainability* 14: 5828. [CrossRef]
- Aslam, Faheem, Paulo Ferreira, Haider Ali, and Sumera Kauser. 2021b. Herding behavior during the COVID-19 pandemic: A comparison between Asian and European stock markets based on intraday multifractality. *Eurasian Economic Review* 12: 333–59. [CrossRef]

- Aslam, Faheem, Paulo Ferreira, Khurram Shahzad Mughal, and Beenish Bashir. 2021c. Intraday volatility spillovers among European financial markets during COVID-19. *International Journal of Financial Studies* 9: 5. [CrossRef]
- Aslam, Faheem, Saqib Aziz, Duc Khuong Nguyen, Khurram S. Mughal, and Maaz Khan. 2020b. On the efficiency of foreign exchange markets in times of the COVID-19 pandemic. *Technological Forecasting and Social Change* 161: 120261. [CrossRef]
- Aslam, Faheem, Tahir Mumtaz Awan, Jabir Hussain Syed, Aisha Kashif, and Mahwish Parveen. 2020c. Sentiments and emotions evoked by news headlines of coronavirus disease (COVID-19) outbreak. *Humanities and Social Sciences Communications* 7: 23. [CrossRef]
- Aslam, Faheem, Yasir Tariq Mohmand, Paulo Ferreira, Bilal Ahmed Memon, Maaz Khan, and Mrestyal Khan. 2020d. Network analysis of global stock markets at the beginning of the coronavirus disease (COVID-19) outbreak. *Borsa Istanbul Review* 20: S49–S61. [CrossRef]
- Barrafrem, Kinga, Daniel Västfjäll, and Gustav Tinghög. 2020. Financial well-being, COVID-19, and the financial better-than-average effect. *Journal of Behavioral and Experimental Finance* 28: 100410. [CrossRef]
- Barro, Robert J., José F. Ursúa, and Joanna Weng. 2020. *The Coronavirus and the Great Influenza Pandemic: Lessons from the “Spanish Flu” for the Coronavirus’s Potential Effects on Mortality and Economic Activity*. Munich: Munich Society for the Promotion of Economic Research—CESifo GmbH.
- Cajueiro, Daniel O., Periklis Gogas, and Benjamin M. Tabak. 2009. Does financial market liberalization increase the degree of market efficiency? The case of the Athens stock exchange. *International Review of Financial Analysis* 18: 50–57. [CrossRef]
- Cao, Guangxi, and Wei Xu. 2016. Nonlinear structure analysis of carbon and energy markets with MFDCCA based on maximum overlap wavelet transform. *Physica A: Statistical Mechanics and Its Applications* 444: 505–23. [CrossRef]
- Caraiani, Petre. 2012. Evidence of multifractality from emerging European stock markets. *PLoS ONE* 7: e40693. [CrossRef] [PubMed]
- Chen, Jilong, Liao Xu, and Hao Xu. 2022. The impact of COVID-19 on commodity options market: Evidence from China. *Economic Modelling* 116: 105998. [CrossRef] [PubMed]
- Consuegra, Meliyara, and Javier Garcia-Verdugo. 2017. Measuring the functional efficiency of agricultural futures markets. *Australian Journal of Agricultural and Resource Economics* 61: 232–46. [CrossRef]
- Corbet, Shaen, Charles Larkin, and Brian Lucey. 2020. The contagion effects of the COVID-19 pandemic: Evidence from gold and cryptocurrencies. *Finance Research Letters* 35: 101554. [CrossRef]
- Czech, Katarzyna, Michał Wielechowski, Pavel Kotyza, Irena Benešová, and Adriana Laputková. 2020. Shaking stability: COVID-19 impact on the Visegrad Group countries’ financial markets. *Sustainability* 12: 6282. [CrossRef]
- Devpura, Neluka, and Paresh Kumar Narayan. 2020. Hourly oil price volatility: The role of COVID-19. *Energy Research Letters* 1: 13683. [CrossRef]
- Domino, Krzysztof. 2011. The use of the Hurst exponent to predict changes in trends on the Warsaw Stock Exchange. *Physica A: Statistical Mechanics and Its Applications* 390: 98–109. [CrossRef]
- Drożdż, Stanisław, Jarosław Kwapien, Paweł Oświęcimka, and Rafał Rak. 2010. The foreign exchange market: Return distributions, multifractality, anomalous multifractality and the Epps effect. *New Journal of Physics* 12: 105003. [CrossRef]
- Drożdż, Stanisław, Paweł Oświęcimka, Andrzej Kulig, Jarosław Kwapien, Katarzyna Bazarnik, Iwona Grabska-Gradzińska, Jan Rybicki, and Marek Stanuszek. 2016. Quantifying origin and character of long-range correlations in narrative texts. *Information Sciences* 331: 32–44. [CrossRef]
- Fama, Eugene F. 1970. Efficient market hypothesis: A review of theory and empirical work. *Journal of Finance* 25: 28–30. [CrossRef]
- Farjah, Ebrahim. 2019. Proposing an efficient wind forecasting agent using adaptive MFDFA. *Journal of Power Technologies* 99: 152.
- Garcin, Matthieu, Jules Klein, and Sana Laaribi. 2020. Estimation of time-varying kernel densities and chronology of the impact of COVID-19 on financial markets. *arXiv* arXiv:2007.09043.
- Goodell, John W. 2020. COVID-19 and finance: Agendas for future research. *Finance Research Letters* 35: 101512. [CrossRef]
- Heng, Yan, Marisa Zansler, and Lisa House. 2020. Orange Juice Consumers Response to the COVID-19 Outbreak: FE1082, 06/2020. *EDIS* 2020: 4.
- Ihlen, Espen A. F., and Beatrix Vereijken. 2013. Multifractal formalisms of human behavior. *Human Movement Science* 32: 633–51. [CrossRef]
- Jafari, G. Reza, Pouria Pedram, and Leila Hedayatifar. 2007. Long-range correlation and multifractality in Bach’s inventions pitches. *Journal of Statistical Mechanics: Theory and Experiment* 2007: P04012. [CrossRef]
- Kantelhardt, Jan W., Stephan A. Zschiegner, Eva Koscielny-Bunde, Shlomo Havlin, Armin Bunde, and H. Eugene Stanley. 2002. Multifractal detrended fluctuation analysis of nonstationary time series. *Physica A: Statistical Mechanics and Its Applications* 316: 87–114. [CrossRef]
- Kelty-Stephen, Damian G. 2017. Threading a multifractal social psychology through within-organism coordination to within-group interactions: A tale of coordination in three acts. *Chaos, Solitons & Fractals* 104: 363–70.
- Kim, Hongseok, Gabjin Oh, and Seunghwan Kim. 2011. Multifractal analysis of the Korean agricultural market. *Physica A: Statistical Mechanics and Its Applications* 390: 4286–92.
- Kim, Irene, and Mohan Venkatachalam. 2011. Are sin stocks paying the price for accounting sins? *Journal of Accounting, Auditing & Finance* 26: 415–42.
- Kumar, Sunil, and Nivedita Deo. 2009. Multifractal properties of the Indian financial market. *Physica A: Statistical Mechanics and Its Applications* 388: 1593–602. [CrossRef]

- Laib, Mohamed, Jean Golay, Luciano Telesca, and Mikhail Kanevski. 2018. Multifractal analysis of the time series of daily means of wind speed in complex regions. *Chaos, Solitons & Fractals* 109: 118–27.
- Laing, Timothy. 2020. The economic impact of the Coronavirus 2019 (COVID-2019): Implications for the mining industry. *The Extractive Industries and Society* 7: 580–82. [CrossRef]
- Li, Zhihui, and Xinsheng Lu. 2011. Multifractal analysis of China's agricultural commodity futures markets. *Energy Procedia* 5: 1920–26. [CrossRef]
- Lu, Ran, and Hongjun Zeng. 2022. VIX and major agricultural future markets: Dynamic linkage and time-frequency relations around the COVID-19 outbreak. *Studies in Economics and Finance. ahead-of-print*. [CrossRef]
- Maier, Benjamin F., and Dirk Brockmann. 2020. Effective containment explains subexponential growth in recent confirmed COVID-19 cases in China. *Science* 368: 742–46. [CrossRef]
- Makowiec, Danuta, Aleksandra Dudkowska, Rafał Gałaska, and Andrzej Rynkiewicz. 2009. Multifractal estimates of monofractality in RR-heart series in power spectrum ranges. *Physica A: Statistical Mechanics and Its Applications* 388: 3486–502. [CrossRef]
- Mandelbrot, Benoit B. 1971. When can price be arbitrated efficiently? A limit to the validity of the random walk and martingale models. *The Review of Economics and Statistics* 53: 225–36. [CrossRef]
- Mandelbrot, Benoit B. 1982. *The Fractal Geometry of Nature*. New York: WH Freeman, vol. 1.
- Mandelbrot, Benoit B. 1997. The variation of the prices of cotton, wheat, and railroad stocks, and of some financial rates. In *Fractals and Scaling in Finance*. Berlin/Heidelberg: Springer, pp. 419–43.
- Mandelbrot, Benoit B., and James R. Wallis. 1969. Robustness of the rescaled range R/S in the measurement of noncyclic long run statistical dependence. *Water Resources Research* 5: 967–88. [CrossRef]
- Mandelbrot, Benoit. 1967. The variation of some other speculative prices. *The Journal of Business* 40: 393–413. [CrossRef]
- McKibbin, Warwick, and David Vines. 2020. Global macroeconomic cooperation in response to the COVID-19 pandemic: A roadmap for the G20 and the IMF. *Oxford Review of Economic Policy* 36: S297–S337. [CrossRef]
- Ming, Wen, Zhengqing Zhou, Hongshan Ai, Huimin Bi, and Yuan Zhong. 2020. COVID-19 and air quality: Evidence from China. *Emerging Markets Finance and Trade* 56: 2422–42. [CrossRef]
- Mishra, Sibanjan. 2019. Testing Martingale Hypothesis Using Variance Ratio Tests: Evidence from High-frequency Data of NCDEX Soya Bean Futures. *Global Business Review* 20: 1407–22. [CrossRef]
- Mnif, Emna, Anis Jarboui, and Khaireddine Mouakhar. 2020. How the cryptocurrency market has performed during COVID 19? A multifractal analysis. *Finance Research Letters* 36: 101647. [CrossRef] [PubMed]
- Muzy, Jean-François, Emmanuel Bacry, Rachel Baile, and Philippe Poggi. 2008. Uncovering latent singularities from multifractal scaling laws in mixed asymptotic regime. Application to turbulence. *EPL (Europhysics Letters)* 82: 60007. [CrossRef]
- Nagy, Zoltan, Peter Mukli, Peter Herman, and Andras Eke. 2017. Decomposing multifractal crossovers. *Frontiers in Physiology* 8: 533. [CrossRef]
- Nicola, Maria, Zaid Alsafi, Catrin Sohrabi, Ahmed Kerwan, Ahmed Al-Jabir, Christos Iosifidis, Maliha Agha, and Riaz Agha. 2020. The socio-economic implications of the coronavirus pandemic (COVID-19): A review. *International Journal of Surgery* 78: 185–93. [CrossRef]
- Njindan Iyke, Bernard. 2020. The disease outbreak channel of exchange rate return predictability: Evidence from COVID-19. *Emerging Markets Finance and Trade* 56: 2277–97. [CrossRef]
- Oh, Gabjin, Seunghwan Kim, and Cheoljun Eom. 2010. Multifractal analysis of Korean stock market. *Journal of the Korean Physical Society* 56: 982–85. [CrossRef]
- Peng, Chung-Kang, Sergey V. Buldyrev, Shlomo Havlin, Michael Simons, H. Eugene Stanley, and Ary L. Goldberger. 1994. Mosaic organization of DNA nucleotides. *Physical Review E* 49: 1685. [CrossRef] [PubMed]
- Pleșoianu, Anita, Alexandru Todea, and Răzvan Căpușan. 2012. The informational efficiency of the Romanian stock market: Evidence from fractal analysis. *Procedia Economics and Finance* 3: 111–18. [CrossRef]
- Podobnik, Boris, and H. Eugene Stanley. 2008. Detrended cross-correlation analysis: A new method for analyzing two nonstationary time series. *Physical Review Letters* 100: 084102. [CrossRef] [PubMed]
- Rizvi, Syed Aun R., Ginanjar Dewandaru, Obiyathulla I Bacha, and Mansur Masih. 2014. An analysis of stock market efficiency: Developed vs Islamic stock markets using MF-DFA. *Physica A: Statistical Mechanics and Its Applications* 407: 86–99. [CrossRef]
- Rosas, Alexandre, Edvaldo Nogueira Jr., and José F Fontanari. 2002. Multifractal analysis of DNA walks and trails. *Physical Review E* 66: 061906. [CrossRef]
- Sadowski, Adam, Zbigniew Galar, Robert Walasek, Grzegorz Zimon, and Per Engelseth. 2021. Big data insight on global mobility during the COVID-19 pandemic lockdown. *Journal of Big Data* 8: 78. [CrossRef]
- Sifat, Imtiaz, Abdul Ghafoor, and Abdollah Ah Mand. 2021. The COVID-19 pandemic and speculation in energy, precious metals, and agricultural futures. *Journal of Behavioral and Experimental Finance* 30: 100498. [CrossRef]
- Sipra, Hassaan, Faheem Aslam, Jabir Hussain Syed, and Tahir Mumtaz Awan. 2021. Investigating the Implications of COVID-19 on PM2. 5 in Pakistan. *Aerosol and Air Quality Research* 21: 200459. [CrossRef]
- Stanley, H. Eugene, and Paul Meakin. 1988. Multifractal phenomena in physics and chemistry. *Nature* 335: 405–09. [CrossRef]
- Stosic, Tatijana, Salman Abarghouei Nejad, and Borko Stosic. 2020. Multifractal analysis of Brazilian agricultural market. *Fractals* 28: 2050076. [CrossRef]

- Štreimikienė, Dalia, Tomas Baležentis, Artiom Volkov, Erika Ribašauskienė, Mangirdas Morkūnas, and Agnė Žičkienė. 2022. Negative effects of covid-19 pandemic on agriculture: Systematic literature review in the frameworks of vulnerability, resilience and risks involved. *Economic Research-Ekonomska Istraživanja* 35: 529–45. [CrossRef]
- Subramaniam, Arvind R., Ilya A. Gruzberg, and Andreas W. W. Ludwig. 2008. Boundary criticality and multifractality at the two-dimensional spin quantum Hall transition. *Physical Review B* 78: 245105. [CrossRef]
- Telesca, Luciano, Vincenzo Lapenna, and Maria Macchiato. 2005a. Multifractal fluctuations in seismic interspike series. *Physica A: Statistical Mechanics and Its Applications* 354: 629–40. [CrossRef]
- Telesca, Luciano, Vincenzo Lapenna, and Maria Macchiato. 2005b. Multifractal fluctuations in earthquake-related geoelectrical signals. *New Journal of Physics* 7: 214. [CrossRef]
- Topcu, Mert, and Omer Serkan Gulal. 2020. The impact of COVID-19 on emerging stock markets. *Finance Research Letters* 36: 101691. [CrossRef]
- Udovichenko, Vladimir, and Peter Strizhak. 2002. Multifractal properties of copper sulfide film formed in self-organizing chemical system. *Theoretical and Experimental Chemistry* 38: 259–62. [CrossRef]
- Umar, Zaghum, Francisco Jareño, and Ana Escribano. 2021. Agricultural commodity markets and oil prices: An analysis of the dynamic return and volatility connectedness. *Resources Policy* 73: 102147. [CrossRef]
- Wang, Jian, Wei Shao, and Junseok Kim. 2020a. Analysis of the impact of COVID-19 on the correlations between crude oil and agricultural futures. *Chaos, Solitons & Fractals* 136: 109896.
- Wang, Yating, Donghao Zhang, Xiaoquan Wang, and Qiuyao Fu. 2020b. How does COVID-19 affect China's insurance market? *Emerging Markets Finance and Trade* 56: 2350–62. [CrossRef]
- Wang, Yudong, and Li Liu. 2010. Is WTI crude oil market becoming weakly efficient over time?: New evidence from multiscale analysis based on detrended fluctuation analysis. *Energy Economics* 32: 987–92. [CrossRef]
- Zhang, Dayong, Min Hu, and Qiang Ji. 2020. Financial markets under the global pandemic of COVID-19. *Finance Research Letters* 36: 101528. [CrossRef] [PubMed]
- Zhang, Yue-Jun, and Shu-Jiao Ma. 2021. Exploring the dynamic price discovery, risk transfer and spillover among INE, WTI and Brent crude oil futures markets: Evidence from the high-frequency data. *International Journal of Finance & Economics* 26: 2414–35.

Article

The Impacts of the Russia–Ukraine Invasion on Global Markets and Commodities: A Dynamic Connectedness among G7 and BRIC Markets

Md. Kausar Alam ¹, Mosab I. Tabash ^{2,*}, Mabruk Billah ³, Sanjeev Kumar ⁴ and Suhaib Anagreh ⁵

- ¹ BRAC Business School, BRAC University, Dhaka 1212, Bangladesh
² College of Business, Al Ain University, Al Ain P.O. Box 64141, United Arab Emirates
³ School of Economics and Finance, Massey University, Auckland 0745, New Zealand
⁴ Faculty of Management Studies, University of Delhi, Delhi 110007, India
⁵ Higher Colleges of Technology, Dubai P.O. Box 25026, United Arab Emirates
* Correspondence: mosab.tabash@aau.ac.ae

Abstract: The conflict between Russia and Ukraine has been causing knock-on effects worldwide. The supply and price of major commodity markets (oil, gas, platinum, gold, and silver) have been greatly impacted. Due to the ongoing conflict, financial markets across the world have experienced a strong dynamic regarding commodities prices. This effect can be considered the biggest change since the occurrence of the financial crisis in the year 2008, which explicitly influenced the oil and gold markets. This study attempts to investigate the impacts of the Russian invasion crisis on the dynamic connectedness among five commodities and the G7 and BRIC (leading stock) markets. We have applied the time-varying parameter vector autoregressive (TVP-VAR) method, which reflects the way spillovers are shaped by various crises periods, and we found extreme connectedness among all commodities and markets (G7 and BRIC). The findings show that gold and silver (commodities) and the United States, Canada, China, and Brazil (stock markets) are the receivers from the rest of the commodities/market's transmitters of shocks during this invasion crisis. This research has policy implications that could be beneficial to commodity and stock investors, and these implications could guide them to make many decisions about investment in such tumultuous situations. Policymakers, institutional investors, bankers, and international organizations are the possible beneficiaries of these policy decisions.

Keywords: Russia and Ukraine conflict; commodities; G7 and BRIC markets; TVP-VAR; connectedness

JEL Classification: G11; G15; H12; J15



Citation: Alam, Md. Kausar, Mosab I. Tabash, Mabruk Billah, Sanjeev Kumar, and Suhaib Anagreh. 2022. The Impacts of the Russia–Ukraine Invasion on Global Markets and Commodities: A Dynamic Connectedness among G7 and BRIC Markets. *Journal of Risk and Financial Management* 15: 352. <https://doi.org/10.3390/jrfm15080352>

Academic Editor: Kentaro Iwatsubo

Received: 29 June 2022

Accepted: 2 August 2022

Published: 8 August 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The conflict between Russia and Ukraine has been causing knock-on effects worldwide. The supply and price of major commodity markets (oil, gas, platinum, gold, and silver) have been greatly impacted.¹ Due to the ongoing conflict, financial markets across the world have experienced a strong dynamic regarding commodities prices. This effect can be considered the biggest change since after the occurrence of the financial crisis in the year 2008, which explicitly influenced the oil and gold markets.² Given this effect, the price of both Brent and West Texas Intermediate (WTI) crude oil has climbed to more than USD 100 per barrel on February 24 while facing the Russian and Ukraine conflict. This invasion has equally changed gas prices, which augmented to USD 3.54 per gallon, and gold prices crossed the figure of USD 1900 per ounce (Liadze et al. 2022).

Accordingly, the prices of commodities are strongly connected with the stock market (Naeem et al. 2022). Therefore, an appropriate connectedness among the five major commodity markets and G7, and BRIC (Brazil, Russia, India and China) markets may be

beneficial for investors in their decision-making processes during the Russian and Ukraine conflict. To the best of our knowledge, this is the first study to examine the rapport between the G7 and BRIC stock and commodity markets before and during the Russia–Ukraine conflict. Thus, the investigation of connectedness among the major commodities and countries will be beneficial for investors and policymakers regarding right and quick decisions for easy investment during the Russian and Ukraine conflict as well as better outcomes by minimizing financial losses.

However, recent studies have found connectedness between the Russia and Ukraine conflict during the short time frame data on key global economies, such as the United States of America, Canada, the United Kingdom, and the European Union (Liadze et al. 2022; Yousaf et al. 2022; Mbah and Wasum 2022). Studies have found negative impacts on the stock market, commodity price, and energy price (Yousaf et al. 2022; Berninger et al. 2022). For example, Yousaf et al. (2022) investigated the conflict between Russia and Ukraine in the G20 and other selected stock markets using the event study approach. They identified that the day of invasion revealed a strong negative impact of this military action on a majority of the stock markets, especially on the Russian market. Tosun and Eshraghi (2022) investigated the financial market reaction to announcements of companies remaining in Russia during the eventful two weeks following the invasion. They found a higher trading volume and selling pressure on remainders, and it was difficult to make any effective decision during the time of political conflict. In general, the Russia–Ukraine war created a challenging economic impact on other countries and on the global economy. Wang et al. (2022) revealed that the total volatility spillover increased from 35% to 85%, exceeding the level seen during the pandemic. The role of commodities changed in both return and volatility spillover systems. Crude oil became a net transmitter of return spillovers, whereas wheat and soybeans became net receivers of return spillovers. Silver, gold, copper, platinum, aluminium, and sugar became net transmitters of volatility. Geopolitical risk Granger caused the spillover indices. High levels of return and volatility spillovers are associated with high levels of geopolitical risk (Wang et al. 2022). The purpose of the current study is to investigate the impacts of the Russian invasion crisis on the financial markets, in particular to identify the main sources of energy market price changes among G7, BRIC and the five commodity markets. According to the recent work by Balcilar et al. (2021), Papathanasiou et al. (2021), and Zhang et al. (2021), the current approach used consists of the time-varying parameter vector autoregressive (TVP-VAR) coming from Antonakakis et al. (2020), which improves the classic technique of Diebold and Yilmaz (2012). Moreover, this method will answer whether these markets' spillovers or connections are higher during the Russia–Ukraine war compared to normal times. We have chosen this methodology because it overcomes restrictions of the basic methodology, as it allows for fluctuations over time and thus provides a more robust estimate. Additionally, the gradation of every roll window width is not an obligatory condition, as roll window analysis is not incorporated, which preserves the use of every available information. Due to the short sample of our paper (1 September 2021–23 February 2022) and during 24 February–24 March 2022, this is a good advantage in case of a conflict between Russia and Ukraine. Moreover, G7 economies represent the developed part of the world and have strategic importance in world GDP, development, trade, investments, and supply chain of commodities (as the largest consumer of the world in PPP) (Waheeduzzaman 2011; Wei et al. 2020; Jiang et al. 2020). Conversely, BRIC markets have played a momentous role in world development, trade, investment, and sectoral cooperation since their inception in 2001 (Iqbal 2021). As a result, BRIC countries in light of other emerging economies (China and India) have emerged as two leading importers (largest consumer base in terms of population) and production hubs of the world, whereas Russia is the principal producer and exporter of energy commodities (Huynh et al. 2020; Shahzad et al. 2019). In the last two decades, the BRIC market group has attracted a large segment of capital inflows, where the highest amount of FDI, FII, and strategic cross border investments are being made (Sauvant 2005; Singhania and Saini 2018; Naeem et al. 2022). Correspondingly, in the last 15 years, the

pace of development has slowed down in developed countries after the global financial crisis (GFC) and the European debt crisis, while BRIC countries have emerged as an engine of world economic growth (Radulescu et al. 2014; Siddiqui 2016).

The empirical analysis discloses that among the other nations, four major economies including the US, Canada, China, and Brazil are the major receivers of losses among G7 and BRICS countries. Similarly, the analysis displays the fact that gold and silver are the receivers from the rest of the commodities/market's transmitters of shocks during this invasion crisis. Our empirical findings will be of interest to market participants and policymakers, as they show that among the five commodities, natural gas remains relatively intact through retransfer mechanisms and can thus form a practical diversification element when added to a portfolio. Similarly, the central banks from these economies should proceed carefully regarding the management of these commodities and should reduce any information asymmetric among the stakeholders of commodities to sustain the market functioning.

The suggesting sections concerning this manuscript are organized as observed: Section 2 describes the review of existing and past literature on the concerned area. Section 3 contains the data and methodology of the paper. Section 4 shows results and discussions. Section 5 concludes the study with some policy implications and limitations.

2. Literature Review

In the past, the invasion of Russia on Ukraine was also considered the most crucial and critical geopolitical disaster, and many worldwide leaders have given their opinions on this crisis.

The current analysis deems to pursue the resource dependency theory in the current perspective. This theory has been utilized by previous literature to see the outcomes in politics. For instance, the analysis of Sprout and Sprout (1957) appeared to not only explore the physical resources, e.g., geography and metals, but also to check the effect of invasions on mental factors including thinking capability and other human reactions. Similarly, another analysis by Pfeffer and Salancik (2003) emphasized the relevant role of scarce and crucial resources, while Beitz (1979) corresponded by stressing at resource fairness that may serve as the root of peace. Advancing the discussion, the study of Reuveny and Barbieri (2014) has explored the relevant impact of war on the utilization of natural resources and has asserted the significant impact on minerals. Selznick (1949) examined the connection between political affairs and enterprises and highlighted the role of political affairs even at the international level on multiple firm-level strategies. Each country owns a specific bundle of resources, e.g., climate, location suitability, fertile land, resources having high demand, and excess availability of common natural resources (Davidson 1980). Given to this, the resource dependency theory supplements a composition to deal with key questions: what are the resources that Russia lacks in terms of quality and quantity? This theory further provides the theoretical background regarding energy sources in Ukraine which are lacking by Russia and urges it for invasion. What are the resources that Russia is interested in acquiring or relocating to their own country? What will be the policy implications of the ongoing war on available resources of Russia and the rest of the world? Hence, the theory facilitates the geographic regions in Ukraine that can be marked as the interest in Russia to be acquired.

The geopolitical risk (GPR) has changed the relationship between European, Russian, and global commodities, where European markets and Russian bonds are collectively transmitting the shocks and affecting returns and volatility in the short and long term (Umar et al. 2022a). Geographical positions of the countries and firms to the war location have implications of returns if countries are located within the boundary of 1000 km, in which it has generated greater negative returns in the four-week time from the war (Federle et al. 2022). Further, during this conflict, results are generated by negative dynamic conditional correlations that USD, JPY, silver, Brent, WTI, and natural gas are found to be a safe haven compared to the Russian rouble to the as indicated (Mohamad 2022).

Additionally, Umar et al. (2022b) found the changes in the behaviour of returns among various financial assets due to GPR even in the normal market conditions, and it is dependent upon the type of market and market situations. Diverse assets depicted different risk patterns in terms of magnitude and timeframe. Bonds and equities have a war impact in the long term, and cryptos have nullified in the short term, while the Swiss franc, gold, silver, green bonds, and oil are the most shock-fighting assets (Bedowska-Sojka et al. 2022). After the Russian invasion, oil was strongly connected with bitcoin, bonds, gold, US dollar, and stocks. Oil also changed its status from a net receiver to a net transmitter of spillovers (Adekoya et al. 2022)

Researchers focused on the relationship between stock markets and energy markets for taking investment decisions and a better understanding of the price fluctuations between the markets (Lin and Su 2020; Peng et al. 2021). The relationship between the two markets have been changed dramatically during the world financial recession, i.e., the global financial crisis (GFC), the great crash of the stock market (GCS), and the European debt crisis (EDC) (Wen et al. 2019, 2020a; Aromi and Clements 2019). The COVID-19 situation also had a significant effect on the global energy markets. In addition, Bouri et al. (2021b) found that US stock, crude oil and gold spillovers seem to intensify during crisis periods. Sharif et al. (2020) outlined that price of oil had a significant effect on US markets and job security, operations of the business, and amenities of mandatory regions were directly impacted in the period of COVID-19. Moreover, Bouri et al. (2021a) found that the dynamic total connectedness across the five assets (gold, crude oil, world equities, currencies, and bonds) was moderate and quite stable during early COVID. Abuzayed et al. (2021) found that bivariate systemic risk contagion between the global stock market and each individual stock market evolved during the sample period and intensified as COVID-19 spread worldwide. Iqbal et al. (2022) found an intensive extreme spillover among the realized volatility of various energy, metals, and agricultural commodities more intensive during the COVID-19 pandemic. As a result, the investors have changed their investment decisions and strategies in stock and energy markets (Mazur et al. 2020; Wen et al. 2020b).

Wars and other natural disasters always hamper economic growth massively. Recently, the two major global economies of Russia and Ukraine have been in the battle and are busy assaulting each other. Both countries are utilizing their military powers to encounter the enemy. This fight has had huge global economic consequences all across the world, as every country is either directly or indirectly globalized in today's time. In addition to such losses, it is further estimated that global GDP will reduce by 1% in the year 2023 due to the globalization effect (World Bank 2022). This loss can be estimated as a USD 1 trillion-dollar reduction in the total GDP of the world. Similarly, the conflict between Ukraine and Russia will add almost 2% to 3% to net inflation across the world (World Bank 2022). In parallel, Ukraine and Russia are major providers of merchandise that include wheat, titanium, corn, etc., on the global stage. Thus, the conflict between both countries can give more to economic complexities regarding the supply of such commodities across the world. Due to the special rebate received by suppliers, the value of such merchandise can move beyond the approximations due to the major chunk and contribution of both states in the global merchandises market. Similarly, this war between Russia and Ukraine can hamper the supply of smartphones, aircraft, and other similar products and thus can intensify the price level of such commodities.

Despite the consequences for other nations, this war can lift the inflation rate to 20 percent in Russia during this year. After COVID-19, this war can prove mounting to more inflation in the Western region of the world. It can be expected that economic growth in the UK can reduce from 0.8 to 4.0 percent in the year 2022 and to 0.5 percent in 2023. Currently, the inflation rate in the UK is 7 percent, which can lower to 5.3 percent excluding the effect of the current war (World Bank 2022). However, the February 2022 outlook report exemplifies that this inflation can go by the rate of 2.7 percent and 2.3 percent in 2023 and 2024, respectively (European Central Bank 2022). The ongoing war between Ukraine and Russia has intensified the other economic issues, e.g., the monetary policy

uncertainty, hampering business confidence, and damaging of overall consumer demand, which was already at the bottom level due to COVID-19-driven price increases. Referring to such damages, it can be further expected that the Russia–Ukraine conflict can increase economic damages on both sides, such as the disruption of trade flows initiates major shortages in the complex food value chain: production, processing, packaging, storage, transportation, and retail sales. In turn, manufacturing will result in excessive logistical costs and high-risk premiums due to missed delivery deadlines and damaged goods (Van Bergeijk 1995). Meanwhile, studies have investigated the connectedness between the commodity price during the COVID-19 period and have found that commodity prices were adversely affected (Mokni et al. 2021; Umar et al. 2021; Iqbal et al. 2022). Wang et al. (2022) studied geopolitical risk and the systemic risk in the commodity markets under the war in Ukraine. They found that a role of commodity changes in both return and volatility spillover systems. Recent studies have found a negative relationship among the global economy, stock market, energy market, commodity price, and resources due to the Russia and Ukraine war (Liadze et al. 2022; Yousaf et al. 2022; Mbah and Wasum 2022; Berninger et al. 2022; Deng et al. 2022). Similarly, investors have an additional penalty due to the ongoing business corporations from the Russia and Ukraine war (Tosun and Eshraghi 2022). Lastly, the world economy is suffering a lot as a result of war crisis (Mbah and Wasum 2022).

Theoretical Review

Even though the prevailing literature provides a sufficient indication of the relevant impact of the Russia–Ukraine 2014 war on the economy, it is uncommon how this ongoing war will affect the efficiency of the commodity market. The existing situation provides credible descriptions of the ongoing conflict between Russia and Ukraine from the past, but fresh evidence is missing. Specifically, several studies are unable to supplement the theoretical background of such conflicts, and thus, a theoretical explanation is missing in the literature. Thus, the current analysis argues the testable hypotheses that fully encompass the role of energy markets and other energy resources, e.g., crude oil in the Russia–Ukraine war. The scholarly evidence on this interesting phenomenon is missing, and the literature has not ascertained the direct role of this conflict on energy markets in both countries (Van de Graaf and Colgan 2017). Belyi (2016) explained some limitations of resource measurements in his study. However, Stulberg (2017) has argued that energy markets and energy act as a tactical curb for Russia, Ukraine, and the European Union. Lee (2017) reveals that the conflict between Ukraine and Russia was aroused due to the historical conflict of gas. Similarly, extracting some more understanding from the analysis of Colgan (2013), it can be further identified that four fundamental paths are playing a fundamental role in the ongoing Russia–Ukraine war. These resources are internal energy markets owned by Ukraine, existing energy resources in Ukraine, Ukraine’s ability to confront the Russian energy dominion in the EU market, transit routes of Ukraine’s gas, and the dependency of the EU and Ukraine on Russian gas (Colgan 2013).

Moreover, recent studies found a negative impact of the Russia and Ukraine war on the global economy, stock market, energy market, commodity price, and resources (Liadze et al. 2022; Berninger et al. 2022; Deng et al. 2022). Tosun and Eshraghi (2022) found that investors have imposed a significant penalty on the remaining firms following the invasion. The review of Mbah and Wasum (2022) revealed that the global economy has begun to feel the impact of this crisis. Inflation, which is already ravaging most global economies, is steadily rising due to the sharp increase in oil, natural gas, and food price shown within a few days of this crisis. Thus, the world economy is experiencing a negative impact on household consumption, increased uncertainty, unpredictable stock swings, supply chain disruptions, bulging utility bills, decreased investment due to political risks, and economic growth impediments. Yousaf et al. (2022), based on a regional analysis, outlined that the European and Asian regions are significantly and adversely affected by this event. Chatziantoniou et al. (2022), in their research, also proved a strong impact of the 2014 war

and other collapses in recent years; more specifically, oil and the Canadian market from G7 are transmitting strong volatility shocks.

3. Data and Methodology

To understand the spillover effects of before (1 September 2021–23 February 2022) and during (24 February–24 March 2022)³ the Russian invasion of Ukraine, we use five major commodity spot prices, namely crude oil (OIL), natural gas (N.GAS), platinum (XPTUSD), silver (XAGUSD), and gold (XAUUSD), and we use the G7 (Canada, France, Germany, Italy, Japan, UK, and US) and BRIC (Brazil, Russia, India, and China) MSCI market indices for the period from 1 September 2021 to 24 March 2022. The chosen countries stand for major advanced and developing economies, affecting global development with their high degrees of commodity needs. Moreover, the data were collected from the Bloomberg database system.

As per Table 1, all the commodities are yielding positive average returns. Except for Canada, all other countries are experiencing a negative average return. Natural gas and crude oil are the most volatile commodities, and Russia has shown the highest volatility followed by the UK and Italy. Here, we may undoubtedly observe the direct impact of the Russian invasion on commodities as well as markets⁴. Here, in Table 1, other than platinum and natural gas, all other commodities including all the sample markets are having negative skewness, which shows that the tail of the distribution is left-skewed and longer or fatter towards the left. Gold, silver, and platinum are out of commodities, and Brazil, the US, and Japan are nearing the standard value of Kurtosis, i.e., 3, which depicts the mesokurtic shape of returns in this distribution. All returns series are stationary at a 1% significance level as per the unit root test of the ADF test (Dickey and Fuller 1979), and the Philips–Perron test (Phillips and Perron 1988).

Table 1. Summary statistics of daily returns of five commodities, G7, and BRIC markets.

Commodities and Stock Markets	Mean	Std. Dev.	Skewness	Kurtosis	JB	ADF	PP
Gold	0.001	0.008	−0.477	3.728	8.34 ***	−26.01 ***	−25.58 ***
Silver	0.001	0.016	−0.395	3.806	7.38 ***	−17.34 ***	−17.40 ***
Platinum	0.001	0.017	0.323	3.685	5.13 ***	−16.04 ***	−16.06 ***
WTI Crude Oil	0.004	0.023	−0.685	8.375	178.17 ***	−19.66 ***	−19.72 ***
Natural Gas	0.001	0.058	0.559	5.960	57.96 ***	−16.76 ***	−16.93 ***
Canada	0.001	0.010	−0.291	4.018	7.97 ***	−23.75 ***	−24.80 ***
France	−0.001	0.014	−0.987	6.573	96.50 ***	−17.83 ***	−18.98 ***
Germany	−0.002	0.015	−0.618	7.370	119.45 ***	−19.66 ***	−18.76 ***
Italy	−0.002	0.024	−2.521	32.564	5209.36 ***	−25.55 ***	−24.57 ***
Japan	−0.001	0.011	−0.325	3.326	3061.89 ***	−13.04 ***	−13.03 ***
UK	−0.002	0.024	−5.770	58.843	18,832.04 ***	−15.27 ***	−16.18 ***
US	−0.001	0.011	−0.258	3.168	1703.51 ***	−11.49 ***	−11.46 ***
Brazil	−0.001	0.017	−0.308	3.173	2369.90 ***	−12.32 ***	−12.35 ***
Russia	−0.009	0.065	−3.918	28.792	4208.21 ***	−18.93 ***	−18.65 ***
India	−0.001	0.016	−2.902	20.200	1908.44 ***	−11.29 ***	−11.29 ***
China	−0.002	0.019	−1.629	14.153	781.84 ***	−15.21 ***	−15.22 ***

Note: The above table illustrates the descriptive statistics for five commodities, G7 and BRIC markets (gold, silver, platinum, WTI Crude Oil, natural gas, Canada, France, Germany, Italy, Japan, UK, US, Brazil, Russia, India, and China). The period was selected daily from 1 September 2021 to 15 March 2022. Moreover, Std. Dev., JB, ADF, and PP represent standard deviations, Jarque-Bera, Augmented Dickey and Fuller, and Phillip and Perron, respectively, with superiors signifying *** $p < 0.01$.

Further, from Figure 1, clear spikes are detected at the end of February and March during the invasion time. Here, all the commodities are presenting positive peaks, while gold, platinum, and crude oil have experienced a greater intensity of volatility (Dodd et al. 2022; Costola and Lorusso 2022). Conversely, all the markets exhibit a downfall, i.e.,

negative volatility has greater impacts than positive shocks supported by many past studies (Dimitriou et al. 2013; Boungou and Yatié 2022; Boubaker et al. 2022).

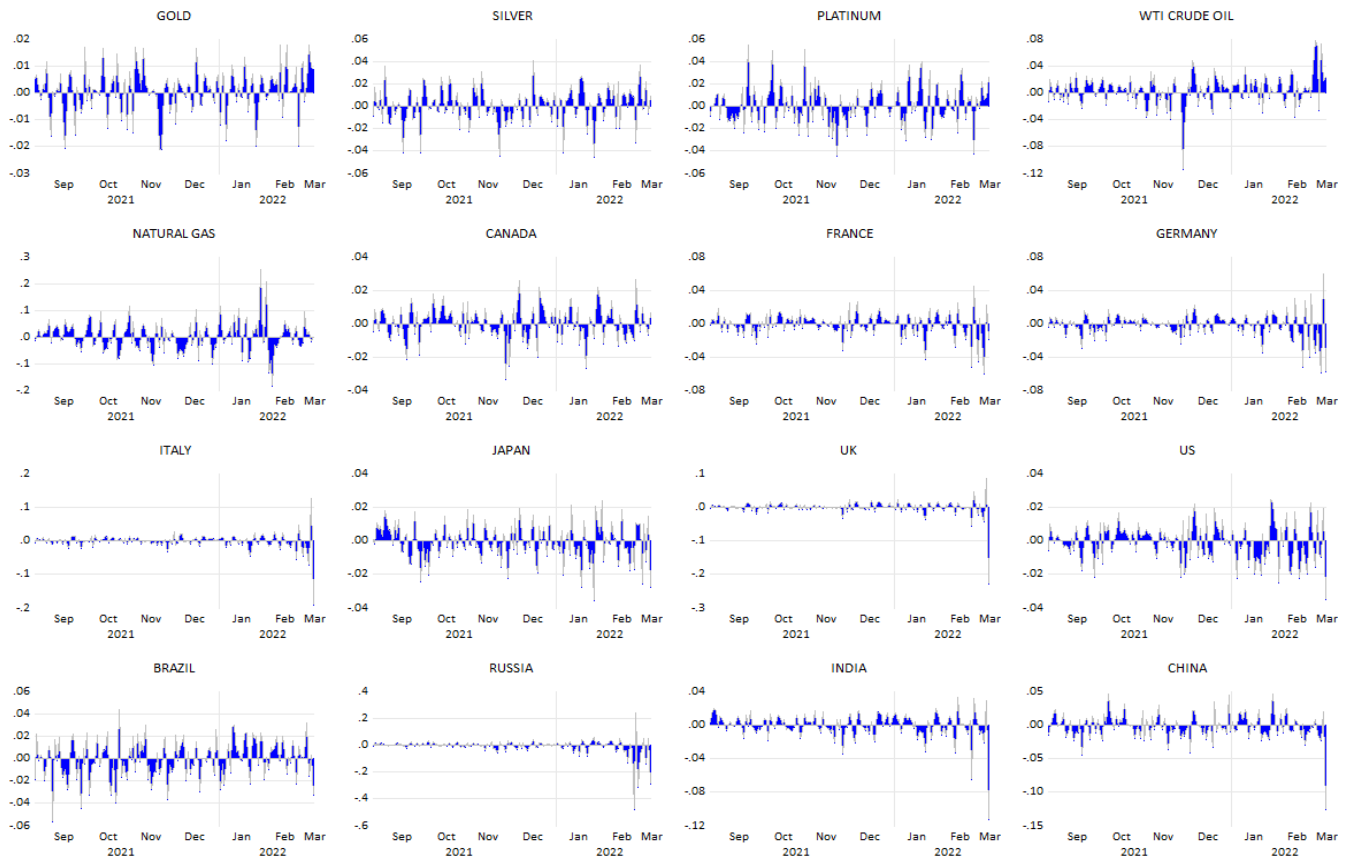


Figure 1. Evolution of five commodities, G7, and BRIC indices from 9 January 2021 to 15 March 2022.

To examine the return spillovers between the five major commodities, G7 and BRIC markets in a time-varying manner, we utilized the TVP-VAR method of Koop and Korobilis (2014) and integrated it using the connectedness method of Diebold and Yilmaz (2014). This particular system enables the variations to differ in time through a Kalman filter evaluation, which depends on the decay elements. By doing this, the TVP-VAR method eliminates the concern of the frequently randomly selected rolling window size, which might cause quite unpredictable or squashed parameters and a lack of important observations (Antonakakis et al. 2018, 2020; Gabauer and Gupta 2018; Korobilis and Yilmaz 2018). This version also provides unique qualities to acknowledge prospective structural breaks and offers considerable factors to acknowledge the connection amongst the factors.

Based upon the Bayesian information criterion (BIC), an autoregressive parameter vector method with time-varying (TVP-VAR) by Antonakakis et al. (2020) is built on the subsequent formula:

$$y_t = A_t Z_{t-1} + \varepsilon_t \quad \varepsilon_t \sim N(0, \Sigma_t) \quad (1)$$

$$vec(A_t) = vec(A_{t-1}) + \zeta_t \quad \zeta_t \sim N(0, \Xi_t) \quad (2)$$

where y_t , Z_{t-1} and ε_t are the $K \times 1$ dimensional vector, and A_t and Σ_t are the $K \times K$ dimensional matrices. $vec(A_t)$ and ζ_t are $K^2 \times 1$ dimensional vectors, whereas Ξ_t is a $K^2 \times K^2$ dimensional matrix. As the dynamic connectedness approach of Diebold and Yilmaz (2012, 2014) rests on the Generalised Forecast Error Variance Decomposition (GFEVD) of

(Koop et al. 1996; Pesaran and Shin 1998), it is required to transform the TVP-VAR to its TVP-VMA representation by the Wold representation theorem:

$$y_t = \sum_{h=0}^{\infty} A_{h,t} \varepsilon_{t-h} \text{ where } A_0 = I_K.$$

The H -step ahead GFEVD models the impact a shock in series j has on series i . This can be formulated as follows:

$$\theta_{ij,t}^g(H) = \frac{\sum_{h=0}^{H-1} (e_i' A_{ht} \Sigma_t e_j)^2}{(e_j' \Sigma_t e_j) \sum_{h=0}^{H-1} (e_i A_t S_t A_t' e_i)} \tag{3}$$

$$\tilde{\theta}_{ij,t}^g(H) = \frac{\theta_{ij,t}^g(H)}{\sum_{k=1}^K \theta_{ij,t}^g(H)} \tag{4}$$

where e_i is a the $K \times 1$ dimensional zero vector with unity on its i th position. As $\theta_{ij,t}^g(H)$ stands for the unscaled GFEVD ($\sum_{j=1}^K \zeta_{ij,t}^g(H) \neq 1$), Diebold and Yilmaz (2009, 2012, 2014) suggested to normalize it by dividing $\theta_{ij,t}^g(H)$ by the row sums to obtain the scaled GFEVD, $\tilde{\theta}_{ij,t}^g(H)$.

The scalable GFEVD is at the core of the connectivity approach and facilitates calculating the total directional connectivity to (from) all indexes from (to) index i . While the total directional connectivity TO describes the effect that index i has on all the others, the total directional connectivity OT describes the impact that all indexes have on index i . These connectivity steps can be calculated by:

$$C_{i \rightarrow j,t}^g(H) = \sum_{j=1, j \neq i}^K \tilde{\theta}_{ji,t}^g(H) \tag{5}$$

$$C_{i \leftarrow j,t}^g(H) = \sum_{j=1, j \neq i}^K \tilde{\theta}_{ij,t}^g(H) \tag{6}$$

Computing the difference between the TO and the FROM total directional connectedness results in the net total directional connectedness of series i :

$$C_{i,t}^g(H) = C_{i \rightarrow j,t}^g(H) - C_{i \leftarrow j,t}^g(H) \tag{7}$$

4. Results and Discussion

This study was conducted on five commodities, G7, and BRIC countries before and during the Russia–Ukraine war. During the invasion crisis, a drastic rise in the prices of commodities, a dramatic fall in the prices of securities, and a huge setback in trade and cross-border investments, more specifically in G-7 and BRIC economies (Wang et al. 2022; Saâdaoui et al. 2022; Orhan 2022) has occurred. This has led to high volatility around the world, especially from the invasion crisis (February 2022-on going). We used daily prices and yield data for five commodities and twelve markets (most developed and developing economies across the world). The data were collected from the Bloomberg database, by applying the formula: $r_{i,t} = \ln(p_{i,t}) - \ln(p_{i,t-1})$, daily return was calculated.

4.1. The Connectedness Network Spillovers

This Russia–Ukraine war has shattered economic activities, trade patterns, market returns and commodities supply chains. We applied the network connectedness of the TVP-VAR method suggested by Koop and Korobilis (2014), which is an advanced version of the traditional Diebold and Yilmaz (2012, 2014) method and estimate for the return spillovers amongst the sample commodities and markets for the period 1 September 2021

to 24 March 2022. Invasion effects can be observed from the results of the invasion on the returns connectedness on commodities and on all sample markets.

From Figure 2a, it can be asserted that prior to the occurrence of the invasion crisis, platinum and natural gas were net recipients of spillovers, and the remaining commodities were net transmitters. It is evident that there is strong connectedness between gold and silver, as both commodities massively influence each other. This description relating to gold and silver has also been stated by (Balli et al. 2019; Naeem et al. 2022; Mbah and Wasum 2022) in their studies. Conversely, the US, China, Japan, and Brazil are the net transmitters with comparatively low intensity, and the rest are recipients. It is quite apparent in the case of capital markets that the UK and other European markets are the most connected markets due to a member of regional economic integration (EU) in the sample countries transmitting the risk/return to each other among European countries. Canada is one of the largest transmitters in the network and is connected to the US, UK, Italy, Germany, and France. The UK is the largest receiver of the spillovers due to major EU countries in the sample data. Before the crisis, Russia, the US, India, China, Japan, and Brazil reflected a lesser connectedness pattern.

Subsequently, an opposite picture is displayed in Figure 2b, where a nest of connections has been presented not only among commodities and capital markets but also within each other, which reflect the consequent effects of the crisis already proven by (Wen et al. 2020a; Bouri et al. 2021a, 2021b; Umar et al. 2022a) in the past, such that commodities were also treated as an alternative investment, more particularly gold and silver. During the invasion crisis, gold and silver are net transmitters, and crude oil, platinum, and natural gas are net recipients. Conversely, most of the capital markets are net transmitters, as they are most affected by the crisis, but only the US, Brazil, China, and Canada are the recipient(s). Conclusively, the ongoing invasion has enormous consequences for sample countries, and it has affected the overall economic positioning of all the sample markets. From the literature, the studies of (Mazur et al. 2020; Bedowska-Sojka et al. 2022; Federle et al. 2022) have also asserted similar effects in the past.

Additionally, a nest is formed among the commodities and markets reflecting high intensity of volatility spillover because risk is being transmitted among them during this GPC. During war, gold and silver among commodities and Japan from markets changed their status from net transmitters to net receivers (Wang et al. 2022). Conversely, natural gas, platinum, and Canada turns net transmitters during the Russian invasion. An interesting observation can be seen that commodities were hardly connected with markets during pre-war time, but huge spillover connectedness is detected during the war (Wen et al. 2020a; Bouri et al. 2021a; Umar et al. 2022a).

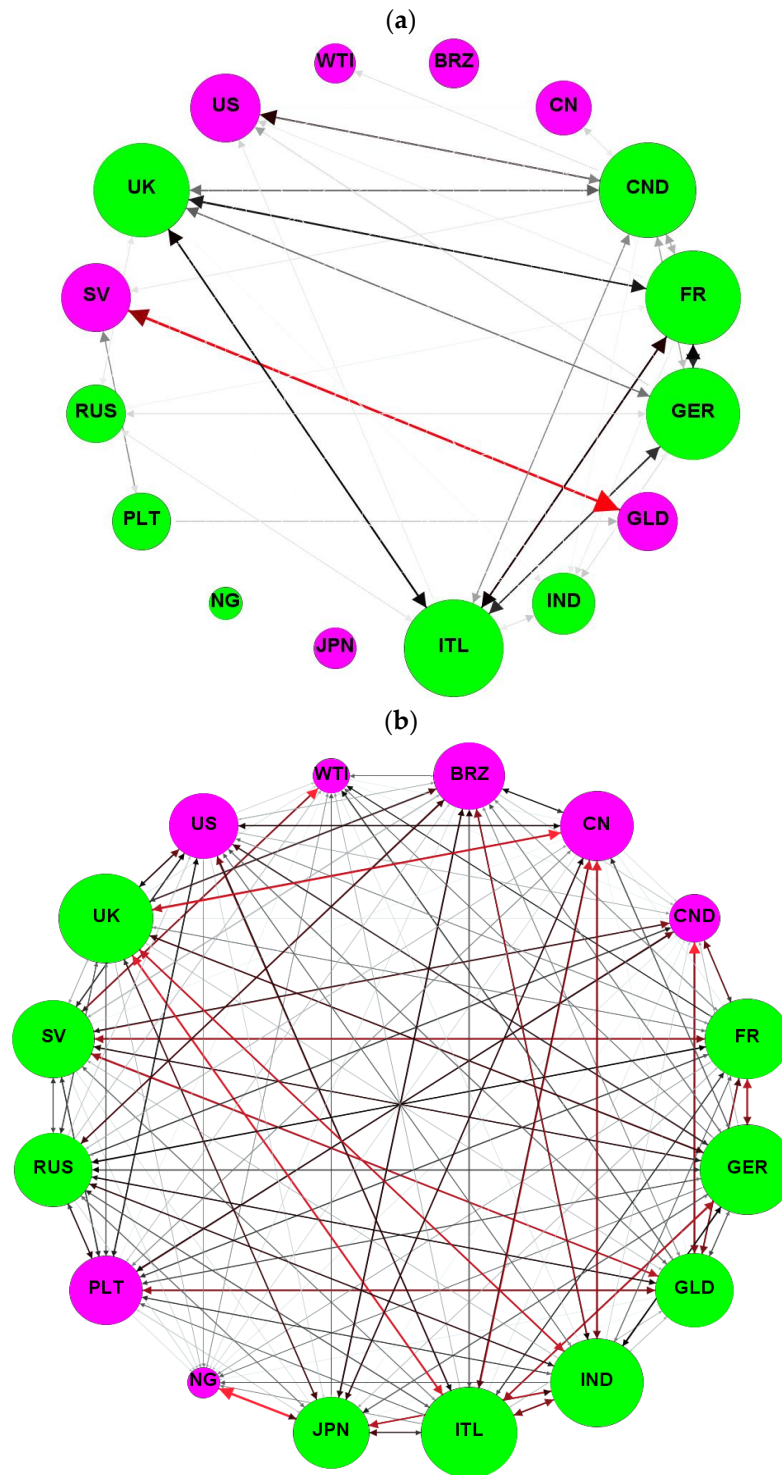


Figure 2. Network connectedness spillovers between the five commodities, G7, and BRIC markets. Additionally, within the network, the size of the node indicates the magnitude of the contribution of every index to the connectivity of the system, while the colour indicates the origin of the connectivity. The size of the node indicates the level of overflow, and the colour determines whether the market is a net sender (green) or a recipient (pink) of spillover. The finite directional layout algorithm determines the position of the vertices, with the number of vectors determining the route of the vertices. The width of the arrow indicates the strength of the multiple gradients, and the colour determines the direction of the gradient from the strongest (red) to the weakest (black). Note: The outcomes are constructed on a first-order TVP-VAR model with a first-order delay length and a 20-level generalized forecast error variance within the estimates. (a) Pre-Russian invasion of Ukraine. (b) During Russian invasion of Ukraine.

4.2. Averaged Total Returns Spillovers

To clarify the effect of ongoing GPC, we have also presented the total time-varying (averaged total returns) spillovers between the five commodities and all the sample countries. In Figure 3, it is shown that before the start of war, the spread of COVID-19 was settling down. The spillover effect was decreasing from its peak level of 86% during the second wave of COVID-19 in the month of September 2021 to around 57% in the month of January 2022. However, this spillover augmented in February due to the sudden start of border tensions between the two companion counterparts. After this, a strong spike in spillover effect was observed that crossed the level of 65%. However, this increasing level stopped and settled at 60%, as the war force was limited and peace talks between the two countries were opened. This again supports the findings of (Adams et al. 2015), which suggest that return spillover collectively increased among all the commodities and markets during war crises (Boungou and Yatié 2022; Chatziantoniou et al. 2022; Umar et al. 2022b). In the process of such uncertain events, even limited diversification opportunities were available due to a high degree of spillovers among all markets and commodities (Wen et al. 2020a; Jiang et al. 2020; Naeem et al. 2022).

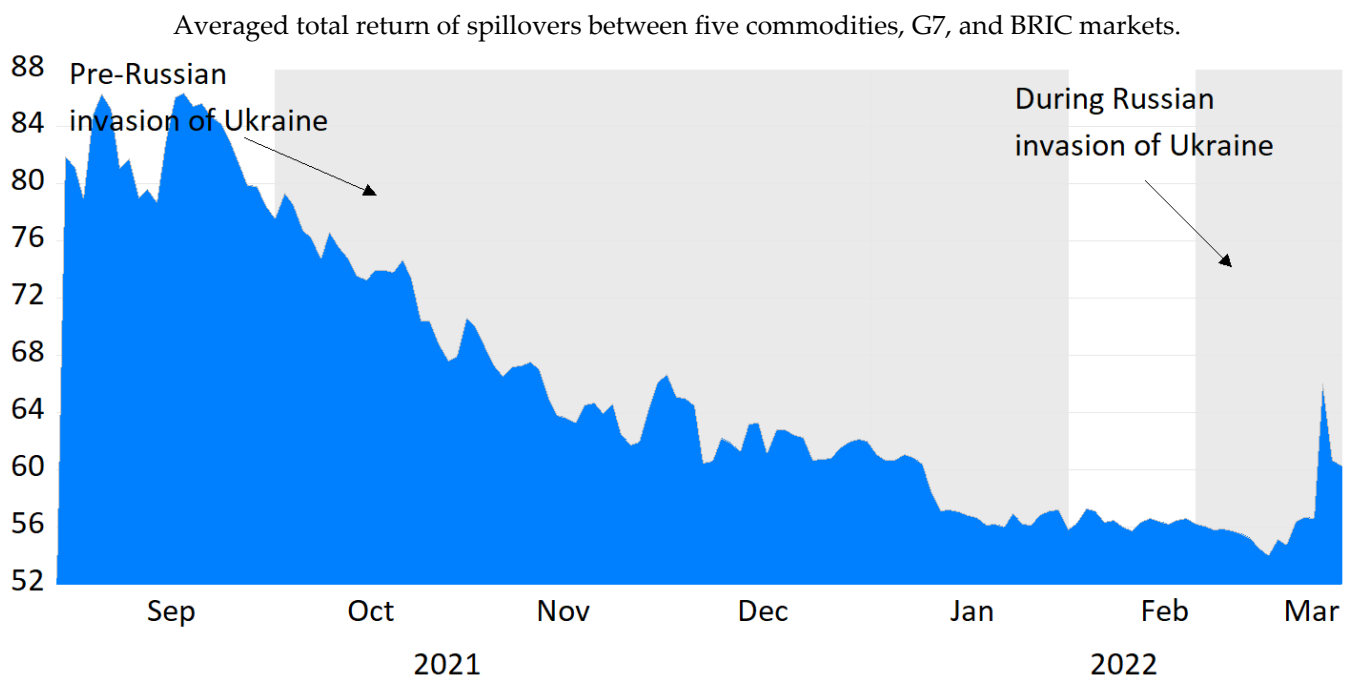


Figure 3. Total time-varying spillovers between five commodities, G7, and BRIC indices. Note: See Figure 2.

4.3. Net Total, “To”, and “From” Return Spillovers

To better understand the spillovers, more specifically during critical periods, we analysed the time-varying behaviour of interconnectedness between commodities and stock markets. Consequently, we also applied the total return spillovers (TO, FROM, NET) as exhibited in Figures 4–6 from all commodities and markets to each commodity and market, respectively. In Figures 4 and 5, total dynamic spillovers to/from each series are displayed and are bidirectional.

Figure 4 shows the spillover transferred to other commodities and markets, where except for natural gas, all other commodities showed a substantial return spillover to other commodities and markets. Platinum, silver, and gold have shown strong spillover variation during the months of February and March even before the invasion started because Russia is one of the largest exporters of these commodities in the world markets⁵. Conversely, almost every market has transmitted return spillover to other markets, and some have reflected spillover effects before the war as well, but post-war peaked spikes can be seen in

each market. Canada seems to be exceptional, as it shows a continuously rising spillover effect since September 2021 due to a slowdown in the economy, but the spillover was further aggravated during the event (Sher 2020). Another important observation is that G7 (except Canada) markets were largely impacted by this war (Federle et al. 2022; Umar et al. 2022b). The US, India, China, and Japan are the largest transmitters to commodities and other markets of the study. This is proven because the US and Japan are one of the largest economies, while India and China are the principal emerging economies in the world.

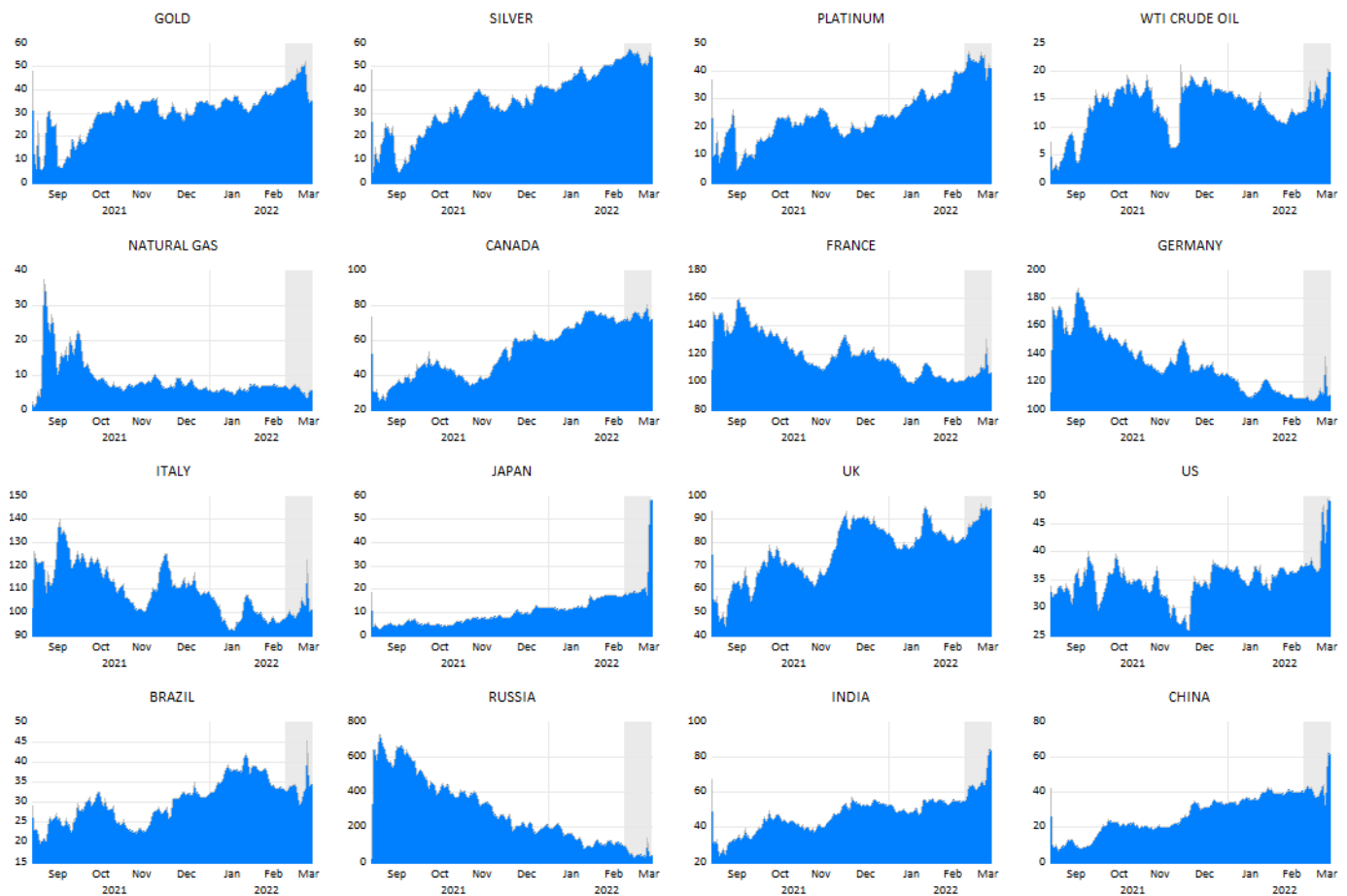


Figure 4. Total return spillovers “TO” others. Note: See Figure 2.

From Figure 5, quite a different image is observed, as WTI crude oil is a prominent recipient of return spillover because EU countries are consuming almost 40% crude oil from Russia (Schiffing and Valantasis Kanellos 2022). Next, platinum, gold, silver, and natural gas (less intensity) are also receiving return spillover from other commodities and markets, but gold and natural gas are experiencing comparatively less spillover effects. In the case of capital markets, other than Canada, all other markets show huge spikes of return spillover from other commodities and markets. Importantly, all European countries were experiencing (receiving) spillover effects not only before the war but also during the war, as they have strong trade ties with both warring countries (Jiang et al. 2020; Berninger et al. 2022; Adekoya et al. 2022). Regarding the BRIC countries, Russia, China, and India are the key players in return spillovers from the commodities and capital markets.

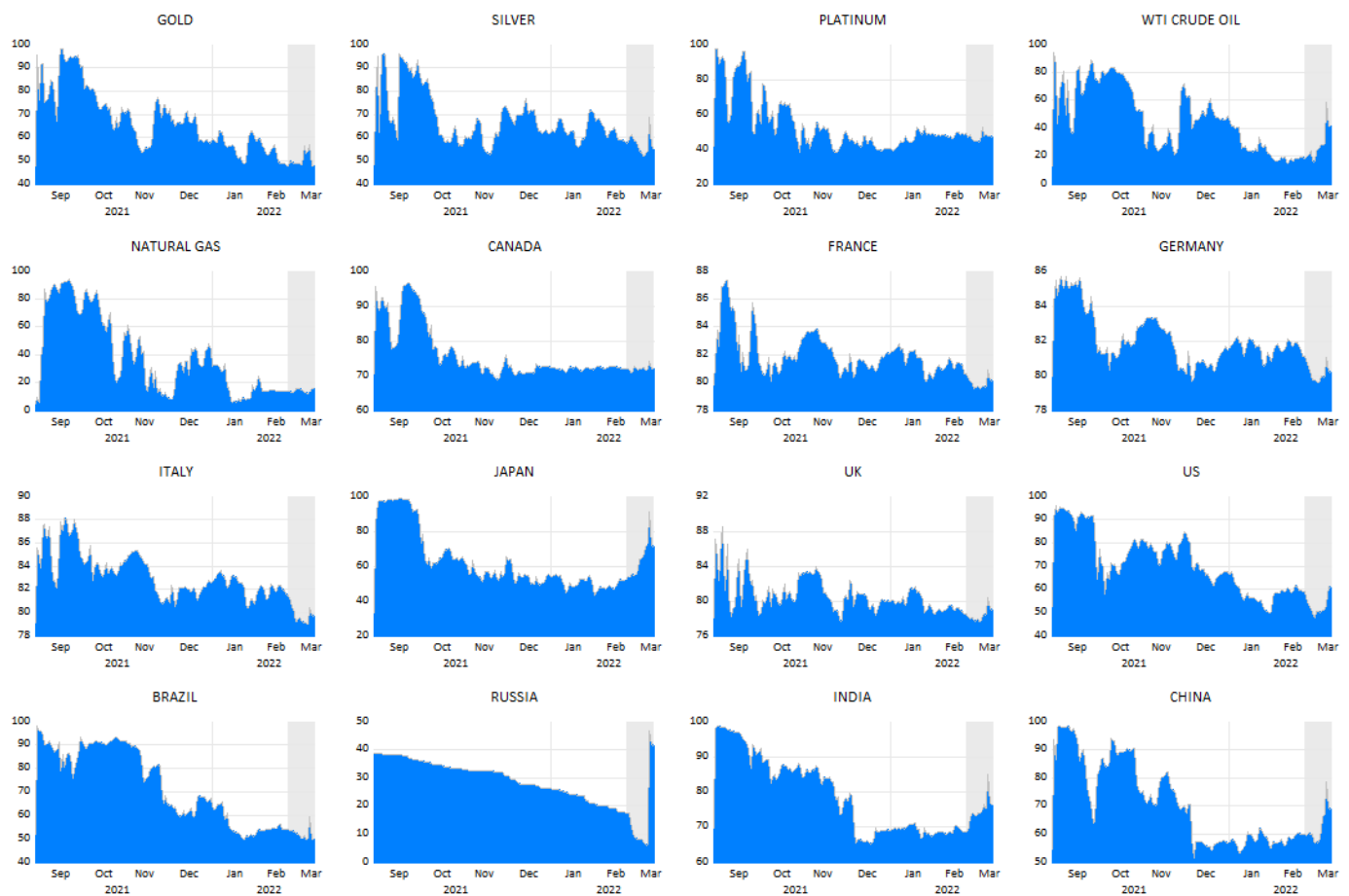


Figure 5. Total return spillovers “FROM” others. Note: See Figure 2.

Additionally, it is observed from Figure 6 that all the commodities are net recipients of return spillovers throughout the sample period, but the quantum is less in the case of natural gas. Crude oil and gold are the most impacted commodities from this invasion crisis, and it is supported by the outcomes from past studies (Billah et al. 2021; Chatziantoniou et al. 2022). Contrariwise, except for the US, China, Japan, and Brazil, all the remaining countries are net transmitters of return spillovers; here, France, Germany, the UK, Italy, and India show rocket spikes. Similar findings were proven by (Adams et al. 2015; Boungou and Yatié 2022; Yousaf et al. 2022; Chatziantoniou et al. 2022) during the war and pandemic crisis situations. All markets are either net recipients or transmitters post-wartime, but India is the only country that was initially a net transmitter and at the end of March, it turned into a net recipient market. This is because the Indian market recovered from the shock nearly to its pre-war level. It is evidently important for the investors, hedgers, and diversifiers from the world to capitalize on this finding on the line of (Mirzaei et al. 2021; Bedowska-Sojka et al. 2022; Mohamad 2022) for international investment diversification.

In past studies, (Yoon et al. 2019; Mensi et al. 2022) have suggested that crisis situations place more emphasis on spillovers, which is somehow matched with this research outcome, i.e., the commodities are total positive transmitters, and at the same time, net total spillover is negative. Hence, all the commodities are net recipients from other commodities and markets. Conversely, our empirical results clearly proved that from the sample G7 and BRIC markets, the US, China, Japan, and Brazil are the net recipients, and the remaining markets have transmitted their losses to other markets and commodities. Thus, special attention should be given to France, Germany, UK, Italy, and India, who have shown rocket spikes (Zhang et al. 2020; Cepoi 2020; Boungou and Yatié 2022; Yousaf et al. 2022), which has proven to be similar to findings in the research of past crisis situations.

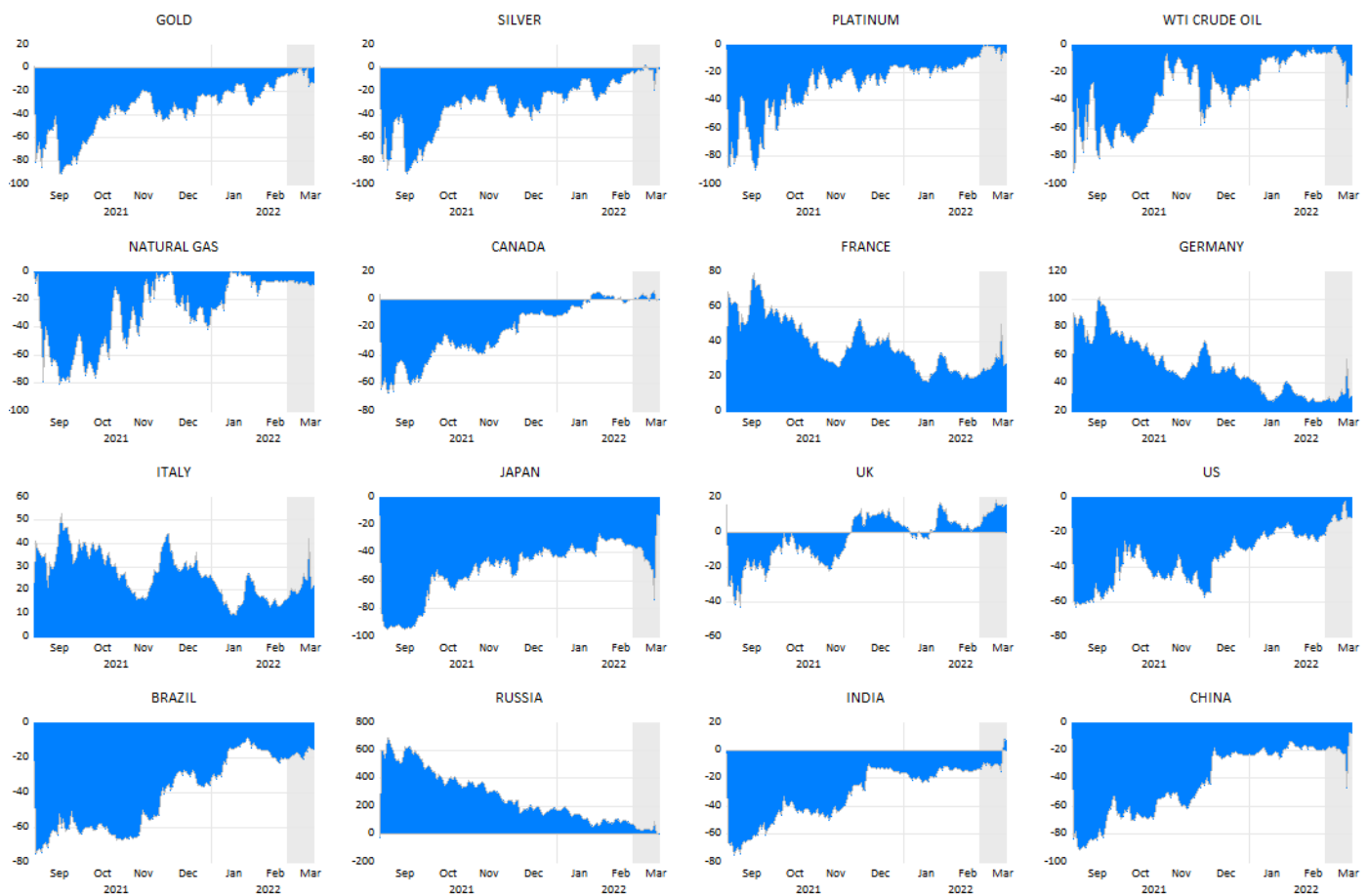


Figure 6. Net total return spillovers between Five commodities, G7, and BRIC markets. Note: See Figure 2.

4.4. Robustness Checks

In this particular section, we performed a few robustness analyses. Along with the TVP-VAR-based connectedness outcomes, we provide 50-day rolling-window VAR and quantile VAR (QVAR) results. Various window sizes happened to be utilized; nevertheless, the 50-observation rolling window revealed close correlations with the TVP-VAR results and is also utilized as a benchmark model in Diebold and Yilmaz (2009, 2012). Given that a VAR model could be determined as an equation-by-equation ordinary least squares (OLS) style, it is a provisional mean-based method and thus is vulnerable to outliers. Suppose we choose each formula by a quantile regression (or the slightest absolute deviation (LAD) regression), in such a case, we concentrate on the conditional median-based computation and can thus probably eliminate the outlier sensitivity issue of the VAR model. Although the dynamics of all three models appear quite comparable, a deeper look discloses that the TVP-VAR model readjusts quicker than its other options, as stressed in Antonakakis et al. (2020) and Korobilis and Yilmaz (2018). This is essential for the forecast of the interconnectedness and thus the risk of the analysed system. This time delay is not too problematic if we only want to track the evolution during the crises. Nevertheless, the outlier sensitivity issue of the VAR model causes inaccurate results, which are more apparent in the Russia–Ukraine war regime.

Figure 7 explains two various sensitivity analyses. Panel A shows the variations in the dynamic total connectedness by readjusting the forecast horizon. We observed that after January 2022, the variations in the measurement enhanced significantly. This could be discussed because the network was more consistent during the Russia–Ukraine war, which showed a boost in its efficiency. Additionally, the variations in the dynamics appeared

to smooth out until the completion of the period, which might lead to the switch of the sample markets back to standard time.

Lastly, Panel B shows the variant of the dynamic connectedness when we enabled the decay factor of the variance–covariance to presume various values. Thus, the decay factor of the VAR coefficient was kept constant at 0.99 because it was unconvincing that the connection throughout variables transforms from one day to another by more than 1%. We discovered that the dot grey area showing the variant of the dynamic connectedness by determining various TVP-VAR requirements did not consist of the dynamic connectedness of the VAR and QVAR values. This marks the time delay issue of the rolling-window models again. The VAR model acted significantly dissimilar to the other two models after January 2022, while the QVAR and the TVP-VAR model shared comparable co-movements.

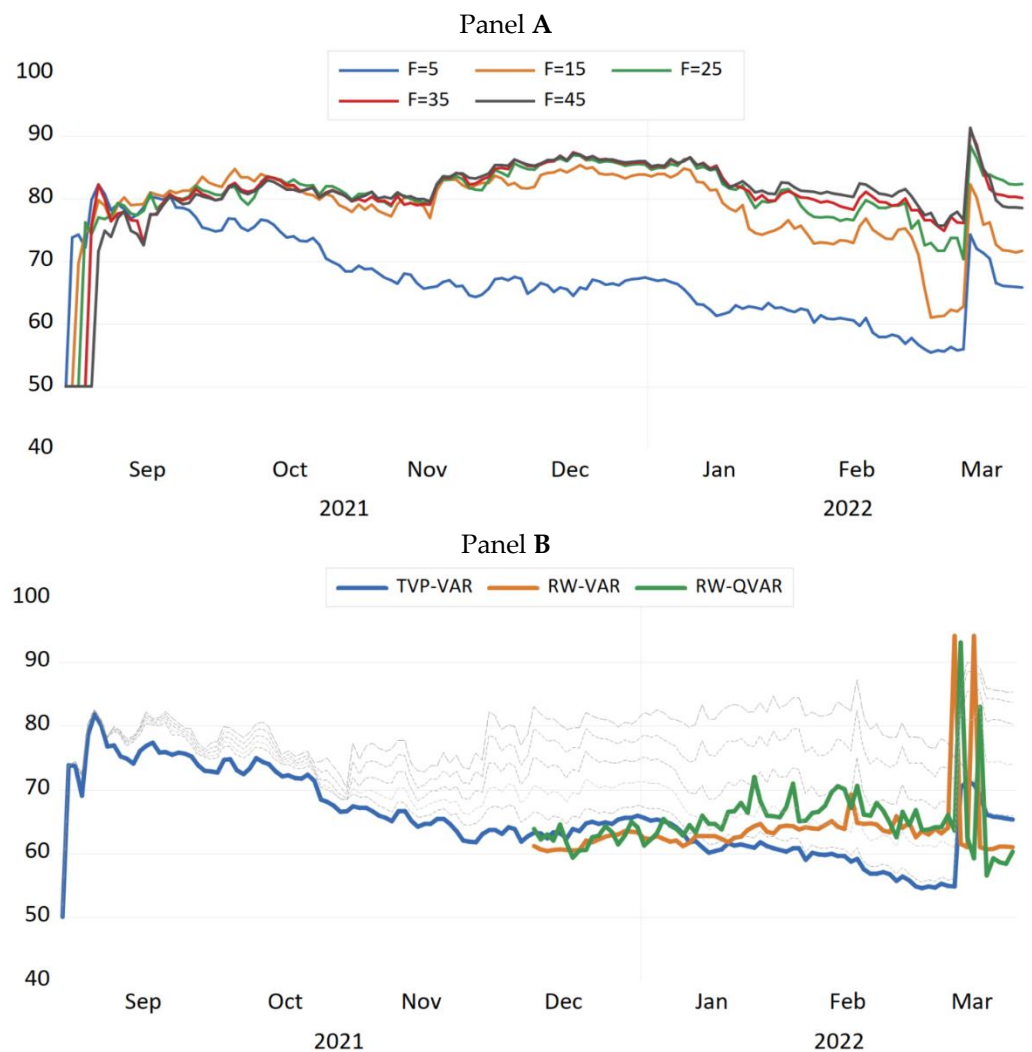


Figure 7. Sensitivity analyses. Note: Panel (A): different forecast horizons are used [5, 15, 25, 35, 45]. Panel (B): $\kappa_1 = [0.95, 0.96, 0.97, 0.98, 0.99]$ and $\kappa_2 = 0.99$. Panel (A): Forecast Horizon Sensitivity Analysis. Panel (B): Decay Factor Sensitivity Analysis.

Our robustness results are also consistent, where we found that after January 2022, the variations in the measurement enhanced significantly and the network was more consistent during the Russia–Ukraine war, which shows a boost in its efficiency. Furthermore, the variations in the dynamics appeared to smooth out until the completion of the period, which might lead to the switch of the sample markets back to standard time. The decay factor of the VAR coefficient was kept constant at 0.99 because it was unconvincing that the connection throughout the variables transforms from one day to another by more than 1%.

5. Conclusions, Policy Implications, and Limitations of the Study

This research investigated the effects of the Russian invasion crisis on the dynamic connectedness between five commodities, G7, and BRIC (leading stock) markets. This study contributed many dimensions to the literature on the spillovers of returns and volatility among sample commodities and markets during GPC caused by the Russian attack on Ukraine. More specifically, return spillovers and volatility behaviour were dissimilar in neighbouring markets (EU) and non-neighbouring markets. This study found that due to this invasion crisis, a very strong connectedness among all commodities and markets (G7 and BRIC) exists. Furthermore, the findings display that gold and silver are the receivers from the rest of the commodities and all the sample markets, whereas platinum, natural gas, silver, and crude oil are the transmitters of shocks during this invasion crisis. Except for the US, Canada, China, and Brazil (recipient), all other countries are net transmitters, where European countries have shown large intensity. Some recent studies found in the literature have also supported the current conclusions of this study, such as (Zhang et al. 2020; Cepoi 2020; Boungou and Yatié 2022; Wang et al. 2022; Yousaf et al. 2022; Chatziantoniou et al. 2022). These studies unveil the phenomenon regarding high market contagion in phases of financial crises in the wake of a huge gain in connectedness in several commodities and financial markets. Particularly, during such a war crisis, global uncertainty has increased and influenced the time-varying connectedness patterns between the commodities and capital markets.

Furthermore, the time-varying net connectedness results express strong responsiveness behaviours among all commodities and capital markets, more specifically among EU markets. This study has policy implications that could be beneficial to commodities and stock investor decisions about investments and hedging in such tumultuous situations. Policymakers, institutional investors, bankers, and international organizations are the potential users to make policy decisions. Geopolitical risk level and connectedness amongst sample commodities and markets could be the guiding force for policymakers to understand the level of systematic risk, in light of these links between commodities and their effect on financial markets, and they could be utilized to prepare strategies to diminish the effects of return spillovers between commodities and stock markets in such crises.

This study was conducted in during a specific period and concluded in a short time, which carries some limitations and will set the path for future research. Due to the paucity of time and dynamicity of the environment, this study has some limitations. First, from the BRICS combination, this study drops South Africa because BRIC countries are the top GDP contributor countries among these major emerging economies, while the South African economy (market) is the least integrated with the rest of the world in terms of trade, investments, markets, and commodities flow⁶ (Waheeduzzaman 2011; Wei et al. 2020; Billah et al. 2021). Second, this study also left out Ukraine and Gulf markets which are the main sources of the commodities, more specifically oil and natural gas. Future research can target these research gaps to give a more robust understanding of this geopolitical crisis. Moreover, further studies can be conducted on sectoral indexes for a wide-ranging investigation of the dynamics of sectoral changes and their risk and returns. However, this study was conducted immediately after the start of the war, and the results are showing short-term consequences; future research might be conducted by taking long-term data sets post-war, which will be useful for diversifiers and hedgers post-war.

Author Contributions: All authors contributed equally to this study. All authors have read and agreed to the published version of the manuscript.

Funding: We have received no funding or any other financial support for the conduct of this research.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data that support the findings of this study is available on request.

Conflicts of Interest: All the authors declare that we have no potential conflict regarding the conduct of research that may interrupt the publication process.

Notes

- ¹ Read more: <https://www.reuters.com/business/energy/is-war-ukraine-impacting-russian-gas-supplies-europe-2022-03-07/> (accessed on 20 June 2022) and <https://www.eia.gov/todayinenergy/detail.php?id=51498> (accessed on 20 June 2022).
- ² Read more: <http://allcoinpedia.com/russia-ukraine-conflict-and-its-impact-on-global-markets/> (accessed on 20 June 2022).
- ³ According to The Guardian website, on 24 February 2022, Russia launched a full-scale invasion of Ukraine: <https://www.jw.org/en/library/series/more-topics/russia-invades-ukraine-bible-meaning-hope/> (accessed on 20 June 2022).
- ⁴ More details at: <https://www.jpmorgan.com/insights/research/russia-ukraine-crisis-market-impact> (accessed on 20 June 2022).
- ⁵ Read more: <https://www.reuters.com/business/energy/is-war-ukraine-impacting-russian-gas-supplies-europe-2022-03-07/> (accessed on 20 June 2022) and <https://www.eia.gov/todayinenergy/detail.php?id=51498> (accessed on 20 June 2022).
- ⁶ Read more at: <https://www.statista.com/topics/1393/bric-countries/#dossierKeyfigures> (accessed on 20 June 2022).

References

- Abuzayed, Bana, EElie Bouri, Nedal Al-Fayoumi, and Naji Jalkh. 2021. Systemic risk spillover across global and country stock markets during the COVID-19 pandemic. *Economic Analysis and Policy* 71: 180–97. [CrossRef]
- Adams, Zeno, Roland Füss, and Felix Schindler. 2015. The sources of risk spillovers among US REITs: Financial characteristics and regional proximity. *Real Estate Economics* 43: 67–100. [CrossRef]
- Adekoya, Oluwasegun Babatunde, Johnson Ayobami Oliyide, OlaOluwa Simon Yaya, and Mamdouh Abdulaziz SalehAl-Faryan. 2022. Does oil connect differently with prominent assets during war? Analysis of intra-day data during the Russia-Ukraine saga. *Resources Policy* 77: 102728. [CrossRef]
- Antonakakis, Nikolaos, David Gabauer, Rangan Gupta, and Vasilios Plakandaras. 2018. Dynamic connectedness of uncertainty across developed economies: A time-varying approach. *Economics Letters* 166: 63–75. [CrossRef]
- Antonakakis, Nikolaos, Ioannis Chatziantoniou, and David Gabauer. 2020. Refined measures of dynamic connectedness based on time-varying parameter vector autoregressions. *Journal of Risk and Financial Management* 13: 23. [CrossRef]
- Aromi, Daniel, and Adam Clements. 2019. Spillovers between the oil sector and the S&P500: The impact of information flow about crude oil. *Energy Economics* 81: 187–96.
- Balcilar, Mehmet, David Gabauer, and Zaghun Umar. 2021. Crude oil future contracts and commodity markets: New evidence from a TVP-VAR extended joint connectedness approach. *Resource Policy* 73: 102219. [CrossRef]
- Balli, Faruk, Muhammad Abubakr Naeem, Syed Jawad Hussain Shahzad, and Anne de Bruin. 2019. Spillover network of commodity uncertainties. *Energy Economics* 81: 914–27. [CrossRef]
- Bedowska-Sojka, Barbara, Ender Demir, and Adam Zarembo. 2022. Hedging Geopolitical Risks with Different Asset Classes: A Focus on the Russian Invasion of Ukraine. *Finance Research Letters* 50: 103192. [CrossRef]
- Beitz, Charles R. 1979. *Political Theory and International Relations*. Cambridge: Cambridge University Press, vol. 13.
- Belyi, Andrei V. 2016. Limitations of resource determinism in international energy studies. *Energy Research & Social Science* 12: 1–4. [CrossRef]
- Berninger, Marc, Florian Kiesel, and Sascha Kolaric. 2022. Should I stay or should I go? Stock Market Reactions to Companies' Decisions in the Wake of the Russia-Ukraine Conflict. April 20. Available online: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4088159 (accessed on 20 June 2022).
- Billah, Mabruk, Sitara Karim, Muhammad Abubakr Naeem, and Samuel Azhgaliyeva Vigne. 2021. Return and volatility spillovers between energy and BRIC markets: Evidence from quantile connectedness. *Research in International Business and Finance* 62: 101680. [CrossRef]
- Boubaker, Sabri, John William Goodell, Dharen Kumar Pandey, and Vineeta Kumari. 2022. Heterogeneous impacts of wars on global equity markets: Evidence from the invasion of Ukraine. *Finance Research Letters* 48: 102934. [CrossRef]
- Boungou, Whelsy, and Alhonita Yatié. 2022. The impact of the Ukraine–Russia war on world stock market returns. *Economics Letters* 215: 110516. [CrossRef]
- Bouri, Elie, Oguzhan Cepni, David Gabauer, and Rangan Gupta. 2021a. Return connectedness across asset classes around the COVID-19 outbreak. *International Review of Financial Analysis* 73: 101646. [CrossRef]
- Bouri, Elie, Xiaojie Lei, Naji Jalkh, Yahua Xu, and Hongwei Zhang. 2021b. Spillovers in higher moments and jumps across US stock and strategic commodity markets. *Resources Policy* 72: 102060. [CrossRef]
- Cepoi, Cosmin Octavian. 2020. Asymmetric dependence between stock market returns and news during COVID19 financial turmoil. *Finance Research Letters* 36: 101658. [CrossRef]
- Chatziantoniou, Ioannis, Christos Floros, and David Gabauer. 2022. Volatility Contagion Between Crude Oil and G7 Stock Markets in the Light of Trade Wars and COVID-19: A TVP-VAR Extended Joint Connectedness Approach. In *Applications in Energy Finance*. Cham: Palgrave Macmillan, pp. 145–68.
- Colgan, Jeff D. 2013. Fueling the fire: Pathways from oil to war. *International Security* 38: 147–80. [CrossRef]

- Costola, Michele, and Marco Lorusso. 2022. Spillovers among energy commodities and the Russian stock market. *Journal of Commodity Markets*, 100249. [CrossRef]
- Davidson, William H. 1980. The location of foreign direct investment activity: Country characteristics and experience effects. *Journal of International Business Studies* 11: 9–22. [CrossRef]
- Deng, Ming, Markus Leippold, Alexander F. Wagner, and Qian Wang. 2022. Stock Prices and the Russia-Ukraine War: Sanctions, Energy and ESG (April 2022). CEPR Discussion Paper No. DP17207. Available online: <https://ssrn.com/abstract=4121382> (accessed on 20 June 2022).
- Dickey, David A., and Wayne A. Fuller. 1979. Distribution of the Estimators for Autoregressive Time Series With a Unit Root. *Journal of the American Statistical Association* 74: 427. [CrossRef]
- Diebold, Francis X., and Kamil Yilmaz. 2009. Measuring financial asset return and volatility spillovers, with application to global equity markets. *The Economic Journal* 119: 158–71. [CrossRef]
- Diebold, Francis X., and Kamil Yilmaz. 2012. Better to give than to receive: Predictive directional measurement of volatility spillovers. *International Journal of Forecasting* 28: 57–66. [CrossRef]
- Diebold, Francis X., and Kamil Yilmaz. 2014. On the network topology of variance decompositions: Measuring the connectedness of financial firms. *Journal of Econometrics* 182: 119–34. [CrossRef]
- Dimitriou, Dimitrios, Dimitris Kenourgios, and Theodore Simos. 2013. Global financial crisis and emerging stock market contagion: A multivariate FIAPARCH–DCC approach. *International Review of Financial Analysis* 30: 46–56. [CrossRef]
- Dodd, Olga, Adrian Fernandez-Perez, and Simon Sosvilla-Rivero. 2022. Currency and commodity return relationship under extreme geopolitical risks: Evidence from the invasion of Ukraine. Available online: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4089030 (accessed on 20 June 2022).
- European Central Bank. 2022. Economic, Financial and Monetary Developments. Economic Bulletin Issue 2. Available online: <https://www.ecb.europa.eu/pub/economic-bulletin/html/eb202202.en.html> (accessed on 8 April 2022).
- Federle, Jonathan, Andre Meier, Gernot J. Müller, and Victor Sehn. 2022. Proximity to War: The stock market response to the Russian invasion of Ukraine. Available online: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4121360 (accessed on 20 June 2022).
- Gabauer, David, and Rangan Gupta. 2018. On the transmission mechanism of country-specific and international economic uncertainty spillovers: Evidence from a TVP-VAR connectedness decomposition approach. *Economics Letters* 171: 63–71. [CrossRef]
- Huynh, Toan Luu Duc, Erik Hille, and Muhammad Ali Nasir. 2020. Diversification in the Age of the 4th Industrial Revolution: The Role of Artificial Intelligence, Green Bonds and Cryptocurrencies. *Technological Forecasting and Social Change* 159: 120188. [CrossRef]
- Iqbal, Badar Alam. 2021. BRICS as a Driver of Global Economic Growth and Development. *Global Journal of Emerging Market Economies* 14: 7–8. [CrossRef]
- Iqbal, Najaf, Elie Bouri, Oksana Grebnevych, and David Roubaud. 2022. Modelling extreme risk spillovers in the commodity markets around crisis periods including COVID19. *Annals of Operations Research*, 1–30. [CrossRef]
- Jiang, Yonghong, Gengyu Tian, and Bin Mo. 2020. Spillover and quantile linkage between oil price shocks and stock returns: New evidence from G7 countries. *Financial Innovation* 6: 1–26. [CrossRef]
- Koop, Gary, and Dimitris Korobilis. 2014. A new index of financial conditions. *European Economic Review* 71: 101–16. [CrossRef]
- Koop, Gary, Mohammad Hashem Pesaran, and Simon M. Potter. 1996. Impulse response analysis in nonlinear multivariate models. *Journal of Econometrics* 74: 119–47. [CrossRef]
- Korobilis, Dimitris, and Kamil Yilmaz. 2018. Measuring Dynamic Connectedness with Large Bayesian VAR Models. Available online: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3099725 (accessed on 20 June 2022).
- Lee, Yusin. 2017. Interdependence, issue importance, and the 2009 Russia-Ukraine gas conflict. *Energy Policy* 102: 199–209. [CrossRef]
- Liadze, Iana, Corrado Macchiarelli, Paul Mortimer-Lee, and Patricia Sanchez Juanino. 2022. The Economic Costs of the Russia-Ukraine Conflict. Available online: <https://www.niesr.ac.uk/wp-content/uploads/2022/03/PP32-Economic-Costs-Russia-Ukraine.pdf> (accessed on 2 March 2022).
- Lin, Boqiang, and Tong Su. 2020. Mapping the oil price-stock market nexus research: A scientometric review. *International Review of Economics and Finance* 67: 133–47. [CrossRef]
- Mazur, Mieszko, Man Dang, and Miguel Vega. 2020. COVID-19 and the march 2020 stock market crash. Evidence from S&P1500. *Finance Research Letters* 38: 101690. [PubMed]
- Mbah, Ruth Endam, and Divine Forcha Wasum. 2022. Russian-Ukraine 2022 War: A Review of the Economic Impact of Russian-Ukraine Crisis on the USA, UK, Canada, and Europe. *Advances in Social Sciences Research Journal* 9: 144–53. [CrossRef]
- Mensi, Walid, Imran Yousaf, Xuan Vinh Vo, and Sang Hoon Kang. 2022. Asymmetric spillover and network connectedness between gold, BRENT oil and EU subsector markets. *Journal of International Financial Markets, Institutions and Money* 76: 101487. [CrossRef]
- Mirzaei, A., M. Saad, and A. Emrouznejad. 2021. Bank Stock Performance During the COVID-19 Crisis: Does Efficiency Explain Why Islamic Banks Fared Relatively Better? Available online: <https://link.springer.com/article/10.1007/s10479-022-04600-y> (accessed on 20 June 2022).
- Mohamad, Azhar. 2022. Safe flight to which haven when Russia invades Ukraine? A 48-hour story. *Economics Letters* 216: 110558. [CrossRef]
- Mokni, Khaled, Mohammed Al-Shboul, and Ata Assaf. 2021. Economic policy uncertainty and dynamic spillover among precious metals under market conditions: Does COVID-19 have any effects? *Resources Policy* 74: 102238. [CrossRef]

- Naeem, Muhammad Abubakr, Linh Pham, Arunachalam Senthilkumar, and Sitara Karim. 2022. Oil shocks and BRIC markets: Evidence from extreme quantile approach. *Energy Economics* 108: 105932. [CrossRef]
- Orhan, Ebru. 2022. The Effects of the Russia-Ukraine War on Global Trade. *Journal of International Trade, Logistics and Law* 8: 141.
- Papathanasiou, Spyros, Dimitrios Vasiliou, Anastasios Magoutas, and Drosos Koutsokostas. 2021. Do hedge and merger arbitrage funds actually hedge? A time-varying volatility spillover approach. *Finance Research Letters* 44: 102088. [CrossRef]
- Peng, Qing, Fenghua Wen, and Xu Gong. 2021. Time-dependent intrinsic correlation analysis of crude oil and the US dollar based on CEEMDAN. *International Journal of Finance and Economics* 26: 834–48. [CrossRef]
- Pesaran, H. Hashem, and Yongcheol Shin. 1998. Generalized impulse response analysis in linear multivariate models. *Economics Letters* 58: 17–29. [CrossRef]
- Pfeffer, Jeffrey, and Gerald R. Salancik. 2003. *The External Control of Organizations: A Resource Dependence Perspective*. Stanford: Stanford University Press.
- Phillips, Peter, and Pierre Perron. 1988. Testing for a unit root in time series regression. *Biometrika* 75: 335–46. [CrossRef]
- Radulescu, Irina Gabriela, Mirela Panaita, and Catalin Voica. 2014. BRICS countries challenge to the world economy new trends. *Procedia Economics and Finance* 8: 605–13. [CrossRef]
- Reuveny, Rafael, and Katherine Barbieri. 2014. On the effect of natural resources on interstate war. *Progress in Physical Geography* 38: 786–806. [CrossRef]
- Saâdaoui, Foued Saâdaoui, Sami Ben Jabeur, and John W. Goodell. 2022. Causality of geopolitical risk on food prices: Considering the Russo–Ukrainian conflict. *Finance Research Letters* 49: 103103. [CrossRef]
- Sauvant, Karl P. 2005. New sources of FDI: The BRICs-outward FDI from Brazil, Russia, India and China. *Journal World Investment & Trade* 6: 639.
- Schiffing, S., and N. Valantasis Kanellos. 2022. Five Essential Commodities That Will Be Hit by War in Ukraine. *The Conversation*. Available online: <https://researchonline.ljmu.ac.uk/id/eprint/16422/1/Five%20essential%20commodities%20that%20will%20be%20hit%20by%20war%20in%20Ukraine.pdf> (accessed on 20 June 2022).
- Selznick, Philip. 1949. *TVA and the Grass Roots: A Study of Politics and Organization*. Berkley: University of California Press, vol. 3.
- Sharif, Arshian, Chaker Aloui, and Yarovaya Larisa. 2020. COVID-19 pandemic, oil prices, stock market, geopolitical risk and policy uncertainty nexus in the US economy: Fresh evidence from the wavelet-based approach. *International Review of Financial Analysis* 70: 101496. [CrossRef]
- Shahzad, Syed Jawad Hussain, Elie Bouri, David Roubaud, Ladislav Kristoufek, and Brian Lucey. 2019. Is Bitcoin a better safe-haven investment than gold and commodities? *International Review of Financial Analysis* 63: 322–30. [CrossRef]
- Sher, Leo. 2020. The impact of the COVID-19 pandemic on suicide rates. *QJM: An International Journal of Medicine* 113: 707–12. [CrossRef] [PubMed]
- Siddiqui, Kalim. 2016. Will the growth of the BRICs cause a shift in the global balance of economic power in the 21st century? *International Journal of Political Economy* 45: 315–38. [CrossRef]
- Singhania, Monika, and Neha Saini. 2018. Determinants of FPI in developed and developing countries. *Global Business Review* 19: 187–213. [CrossRef]
- Sprout, Harold, and Margaret Sprout. 1957. Environmental factors in the study of international politics. *Journal of Conflict Resolution* 1: 309–28. [CrossRef]
- Stulberg, Adam N. 2017. Natural gas and the Russia-Ukraine crisis: Strategic restraint and the emerging Europe-Eurasia gas network. *Energy Research & Social Science* 24: 71–85. [CrossRef]
- Tosun, Onur Kemal, and Arman Eshraghi. 2022. Corporate decisions in times of war: Evidence from the Russia-Ukraine Conflict. *Finance Research Letters* 48: 102920. [CrossRef]
- Umar, Zaghum, Mariya Gubareva, and Tamara Teplova. 2021. The impact of Covid-19 on commodity markets volatility: Analyzing time-frequency relations between commodity prices and coronavirus panic levels. *Resources Policy* 73: 102164. [CrossRef]
- Umar, Zaghum, Ahmed Bossman, Sun-Yong Choi, and Tamara Teplova. 2022a. Does geopolitical risk matter for global asset returns? Evidence from quantile-on-quantile regression. *Finance Research Letters* 48: 102991. [CrossRef]
- Umar, Zaghum, Onur Polat, Sun-Yong Choi, and Tamara Teplova. 2022b. The impact of the Russia-Ukraine conflict on the connectedness of financial markets. *Finance Research Letters* 48: 102976. [CrossRef]
- Van Bergeijk, Peter. 1995. The impact of economic sanctions in the 1990s. *The World Economy* 18: 443–55. [CrossRef]
- Van de Graaf, Thijs, and Jeff D. Colgan. 2017. Russian gas games or well-oiled conflict? Energy security and the 2014 Ukraine crisis. *Energy Research & Social Science* 24: 59–64. [CrossRef]
- Waheeduzzaman, A. N. M. 2011. Competitiveness and convergence in G7 and emerging markets. *Competitiveness Review: An International Business Journal* 21: 110–28. [CrossRef]
- Wang, Yihan, Elie Bouri, Zeeshan Fareed, and Yuhui Dai. 2022. Geopolitical risk and the systemic risk in the commodity markets under the war in Ukraine. *Finance Research Letters* 49: 103066. [CrossRef]
- Wei, Hua, Syed Kumail Abbas Rizvi, Ferhana Ahmad, and Yuchen Zhang. 2020. Resource cursed or resource blessed? The role of investment and energy prices in G7 countries. *Resources Policy* 67: 101663. [CrossRef]
- Wen, Fenghua, Longhao Xu, Guangda Ouyang, and Gang Kou. 2019. Retail investor attention and stock price crash risk: Evidence from China. *International Review of Financial Analysis* 65: 101376. [CrossRef]

- Wen, Fenghua, Yujie Yuan, and Wei-Xing Zhou. 2020a. Cross-shareholding networks and stock price synchronicity: Evidence from China. *International Journal of Finance & Economics* 26: 914–48.
- Wen, Fenghua, Cui Li, Han Sha, and Liuguo Shao. 2020b. How does economic policy uncertainty affect corporate risk-taking? Evidence from China. *Finance Research Letters* 41: 101840. [CrossRef]
- World Bank. 2022. Global Growth to Slow through 2023, Adding to Risk of 'Hard Landing' in Developing Economies. Report. Available online: <https://www.worldbank.org/en/news/press-release/2022/01/11/global-recovery-economics-debt-commodity-inequality> (accessed on 20 June 2022).
- Yoon, Seong-Min, Md Al Mamun, Gazi Salah Uddin, and Sang Hoon Kang. 2019. Network connectedness and net spillover between financial and commodity markets. *The North American Journal of Economics and Finance* 48: 801–18. [CrossRef]
- Yousaf, Imran, Ritesh Patel, and Larisa Yarovaya. 2022. The Reaction of G20+ Stock Markets to the Russia-Ukraine Conflict 'Black-Swan' Event: Evidence from Event Study Approach. March 29. Available online: <https://ssrn.com/abstract=4069555> (accessed on 20 June 2022).
- Zhang, Dayong, Min Hu, and Qiang Ji. 2020. Financial markets under the global pandemic of COVID-19. *Finance Research Letters* 36: 101528. [CrossRef]
- Zhang, Hua, Jinyu Chen, and Liuguo Shao. 2021. Dynamic spillovers between energy and stock markets and their implications in the context of COVID-19. *International Review of Financial Analysis* 77: 101828. [CrossRef]

Article

A Cyclical Phenomenon among Stock & Commodity Markets

Hector O. Zapata *, Junior E. Betanco, Maria Bampasidou and Michael Deliberto

Department of Agricultural Economics & Agribusiness, Louisiana State University, LSU Agricultural Center, Baton Rouge, LA 70803, USA; jbetan4@lsu.edu (J.E.B.); mbampa1@lsu.edu (M.B.); mdeliberto@agcenter.lsu.edu (M.D.)

* Correspondence: hzapata@lsu.edu

Abstract: Considerable studies have examined the relationship between commodity markets and stock markets. This paper studies the cyclical relationship between commodity markets and stock markets with implications for investing based on index relationships. Stock markets are represented by the U.S. S&P 500 index and aggregate commodity markets by the U.S. producer price index (PPI). Tradeable market indexes readily available to investors, namely the S&P GSCI Index and the Bloomberg Commodity Index (BCOM), are also studied. An optimal bandpass filter is used to estimate the cyclical component using a pricing-performance measure of the S&P 500 relative to the PPI, based on annual data from 1871 to 2022. The S&P GSCI and the BCOM indexes are also used to test the robustness of the findings. The impacts of the financial crisis of 2008 and the coronavirus pandemic are also assessed. The overriding conclusion of the study is that a cyclical relationship exists between stock markets and commodity markets for both aggregate and tradeable indexes which can last, from peak to peak, approximately 31 years. Measuring returns and risks in a manner consistent with these cycles can shed new light on the usefulness of commodity investing via tradeable indexes in seeking efficient portfolios.

Keywords: commodity; stock; markets; cycles; investing; risk; returns



Citation: Zapata, Hector O., Junior E. Betanco, Maria Bampasidou, and Michael Deliberto. 2023. A Cyclical Phenomenon among Stock & Commodity Markets. *Journal of Risk and Financial Management* 16: 320. <https://doi.org/10.3390/jrfm16070320>

Academic Editor: Kentaro Iwatsubo

Received: 31 May 2023

Revised: 30 June 2023

Accepted: 1 July 2023

Published: 4 July 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Commodity markets go hand in hand with the history of mankind. In the early part of this history, commodities were used primarily as sources of human food, feed, and for rudimentary toolmaking. Access to commodity markets, either through cash transactions (spot) or futures markets (derivatives), has now become widespread, and the monetary value of commodities in financial markets today is worth trillions of dollars. Surprisingly, however, it has been only over the past two decades that investment interest in commodity markets by market observers, institutional investors, and academic researchers has increased (e.g., Bannister and Forward 2002; Beenen 2005; Rogers 2004; Fabozzi et al. 2008; Zapata et al. 2012; Rouwenhorst and Tang 2012; Bhardwaj et al. 2015; Irwin et al. 2020; Hernandez et al. 2021; Billah et al. 2023, among many others). One factor contributing to the paucity of adopting commodities in investor's portfolios has been driven by misconceptions surrounding commodities, such as the fact that commodities are riskier than equities, there is no relationship between equity and commodity returns, and commodity risk premiums are different from those of equities (facts which the above literature has helped demystify). While many of the unique aspects of commodity markets continue to be issues of research interest, one phenomenon of a recurrent nature that has gained interest in the most recent investment literature is that of the relationship between stock and commodities prices that in the words of Prescott (1986) can be phrased as a cyclical phenomenon. Zapata et al., using an econometric model of cycles, were the first to quantify this relationship following the work of Banister and Forward using aggregated time series data. Rogers, known as a commodity investing guru in the business news media, calls commodities "hot" and "the world's best market," Irwin et al. find that the real-time performance of commodity

futures investing has been disappointing. When considering investing in commodities, it must be noted that commodity markets have characteristics that separate them from traditional equities. First, the demand for most commodity products is inelastic (Tomek and Robinson 1991), and with a growing population, the demand for commodities continues to increase (Evans and Lewis 2005). Second, commodities are positively correlated with inflation, particularly unexpected inflation (Gorton and Rouwenhorst 2006), and as a result, commodities may serve as an inflation hedge. Third, commodities can provide equity-like returns; Bodie and Rosansky (1980), for instance, found that using commodity and stock returns from 1950–1976, the mean returns on commodity futures were similar to those of stocks. Likewise, Greer (2000), after analyzing the returns of stocks and commodities during the period 1970 to 1999, concluded that the returns in both assets were similar. Fourth, commodity price behavior is strongly linked to fluctuations in commodity-specific market fundamentals (supply, demand, biology, and natural factors) which may not affect traditional equities. These characteristics make commodities worth considering in the search for efficient portfolios.

While commodity markets are unique, their relationship to broad financial markets has gained recent interest in the finance literature, and it has been observed that their relationship tends to fluctuate with the different phases of the business cycle (Bhardwaj and Dunsby 2013). Naturally, commodity markets have exposure to natural and biological disasters, geopolitical events, financial crises, etc., that can impact the nature of their relationship to equity markets and, as such, impact the returns and risks trade-offs resulting from each asset's price dynamics.

The purpose of this article is to investigate the historical pricing performance of commodity markets relative to traditional equities, and consistent with Zapata et al., pricing performance is used to measure the relative strength (RS) of the relationship between the S&P 500 index and an aggregate measure of commodity prices, the U.S. producer price index (PPI), over the 1871–2022 period. Unique to this article, however, is the introduction of tradeable indexes, in particular, the first major investible commodity index (the S&P GSCI), and the Bloomberg Commodity Index (BCOM)¹, which have data available starting in 1960 and 1970, respectively, and that later became available to the investment community. The article further investigates the impact of two major crises, the financial crisis of 2008–2009 and the COVID-19 pandemic, on the RS relationship and historical cyclical pattern between commodity markets and the broad equity markets prior to conducting the cyclical analysis. This is carried out in order to ascertain that the cycles approach presented later in this paper does not require special modeling adaptations. It turns out that the financial crisis of 2008–2009 had a longer recessionary impact on the RS performance while the impact of the COVID-19 pandemic on the RS measure was more transitory. However, both impacts did not appreciably change the long-term trend in the RS but it took much longer to return to the trend after the financial crisis. The identified cyclical pattern that emerges is used to further measure returns and risks to commodity investing in long-investment horizons that are consistent with the RS pattern. The article focuses strictly on commodity markets using two tradeable commodity indexes, one of which is the BCOM, is based on futures markets. The overriding conclusion of our research is that the cyclical relationship between commodities and traditional equities remains very strong, lasting approximately 31 years from peak to peak, in consistency with published research. Unique to this article is the conclusion that the aggregate relationship between commodities and equities is maintained at a much lower level of disaggregation using tradeable indexes that have been available to investors for the past several decades. The finding is relevant to portfolio analysis that stocks and commodities alternate in pricing leadership in a synchronized “bulls” versus “bears” interplay that lasts almost two decades for each market.

2. Review of Literature

The cyclical phenomenon in commodity prices (agriculture-crops and livestock, industrial and precious metals, and energy) in relation to equity markets has been of increasing

interest to market observers, investors, and academic researchers over the past two decades. A Web of Science (WoS) search of the literature using the keywords “commodity cycles and agriculture” and the “All Fields” option in the database, over the period January 2002 to November 2022, and conducted in early December 2022, revealed an increasing growth in publications. We added the word “agriculture” to the search to ascertain the inclusion of agricultural commodities. Traditionally, studies of commodity prices and volatility have focused on metals and energy, but agricultural commodities (spot and futures markets) have been of increasing investor interest. The results generated a cross-sectional literature including energy fuels, metals, environmental sciences, and agriculture multidisciplinary. The number of publications per year has increased from 4 in 2002 to 34 by November 2022; similarly, total citations over the same period increased from 4 in 2002 to more than 1100. Selected and well-cited articles from these results are reviewed.

In addressing the inquiry of whether commodity-serving companies deserve investor capital, Bannister and Forward (2002) graphically examined the U.S. equity market index performance relative to the commodity market index and observed that stocks and commodities have alternated leadership in a regular cyclical wave averaging 18 years. They observed that when commodity prices declined, stocks rose 11.6% per year (stock leadership), but stocks increased only 3.4% per year during inflation (commodity leadership), thus lending support to the existence of a cyclical phenomenon on the relationship between equity markets and commodity returns. Banister and Forward did not build an econometric model to estimate an approximate optimal bandpass filter for the relative market index performance.

Radetzki (2006) identified three major commodity booms since the second world war, namely 1950–51 (in response to the Korean war), 1973–74 (intensified by harvest failures and high energy costs), and 2004–onwards (which coincided with the explosive growth in China and India and their demand for raw materials). Radetzki’s results highlight that not all commodities responded with the same intensity during these three boom periods. It was found that the second boom was the most powerful in aggregate terms and was driven by the strong increase in energy prices in 1973 and 1974. Similarly, agricultural raw material prices dominated during the first boom while metals and mineral prices had the sharpest increase in the third boom. Radetzki did not attempt to discuss the implications of booms in commodity markets relative to equity markets as in Banister and Forward. However, it is noted from illustrations in Banister and Forward that the performance of stocks was superior to that of commodities from 1950 to the late 1960s, with leadership returning to commodities during the 1970s and early 1980s. Radetzki asserted that demand shocks tend to trigger commodity booms, and increases in global growth in GDP and industrial production preceded or marked the beginning of the three commodity booms. How commodity markets respond to macroeconomic performance differs from responses in traditional equity markets, and such responses can drive equity and commodity market prices in waves that last several years; this suggests that the use of the relative-strength performance measure of equity market prices to commodity market prices is a reasonable indicator of when returns in one market lead the other. An important consideration in the above two studies is that good economic performance does not always result in booming commodity market prices. Other factors unique to commodity markets such as tight production capacity and relatively small inventories emerge after long periods of low commodity prices which discourage investments in capacity expansion.

Beenen (2005) points out that institutional investor interest in commodities reflects powerful cyclical and structural forces benefiting commodity markets but also reflecting the deterioration of equity and fixed income returns of the time (compared to the 1990s). To satisfy growing investor demand, Deutsche Bank developed an Investor Guide to Commodities that put commodities into a distinct asset class, and as Rogers (2004) stated the bull market in commodities that was underway which was noted to last about 30 years or so. The guide also tried to dispel the myth about commodities that, for most people, commodities imply an elevated level of risk; yet investing in commodities does imply risks

that may differ from the risk of investing in stocks and bonds, and that at certain times in the business cycle commodities have been a much better investment than most alternatives.

One of the first analyses on the returns to investing in commodity futures when compared to investing in stocks and bonds is Gorton and Rouwenhorst (2006). They constructed an equally weighted index of commodity futures monthly returns for July 1959 to March 2004 and found that fully-collateralized commodity futures have historically produced similar returns and Sharpe ratios as equities (this finding is consistent with that of Bodie and Rosansky (1980)'s data from 1949–1976). One finding in Gorton et al.'s analysis was that the historical risk premium of commodity futures is about equal to the risk premium of stocks and is more than double the risk premium of bonds, but also find that commodity futures returns are negatively correlated with those of equities and bonds. They emphasize that, to a large extent, the negative correlation is driven by different behavior over the business cycle and a positive correlation between commodity futures and inflation. While this paper did not have a large cross-section of commodities, it served to illustrate the portfolio diversification value of commodities relative to stocks and bonds. While the behavior over the business cycle is mentioned, no attempt was made to estimate the cycle in commodities relative to equities.

The correlation of commodities to stocks and bonds using ex-post business cycle conditions has been explored in the financial literature. Kat and Oomen (2006) reassess the correlation between commodities and stocks and bonds by arguing that commodities are fundamentally different from financial assets and that there are at least two reasons for a negative correlation. First, commodity prices are the result of current market conditions (rather than long-term prospects) and tie this to the business cycle by suggesting that commodities are likely to do best during the expansion phase and their worst during the recession phase. Second, commodities are more exposed to event shocks (e.g., supply cuts by OPEC, bad weather for crops, or strikes for mining) which can drive commodity prices up and produce positive returns for investors who are long in such commodities; the increased cost of raw materials, however, can put pressure on stocks. Using correlation coefficients and Kendall's Tau between the daily excess returns on 27 commodity futures and the daily excess returns on the Dow Jones Industrial Average (DJIA), and using the whole period as well as different phases of the business cycle, they found that daily commodity futures returns are only weakly correlated with equity returns, little persistence in the correlation dynamics, and that the correlation with equity can vary over the different phases of the business cycles using NBER dating methodology, especially for energy and metals. Contradicting Gorton and Rouwenhorst, Kat and Oomen did not find that the diversification benefits of commodity futures tend to be larger at longer horizons. Kat and Oomen did not estimate the pricing performance of stocks and bonds relative to commodity markets using cycle methodologies but used business cycles based on NBER dating.

The factors that contribute to an increased interest in commodity markets by financial investors were studied by Domanski and Heath (2007). Commodity prices of energy, precious metals, base materials, and agriculture, measured by the Goldman Sachs commodity index (GSCI), and derivative activity, measured by the number of contracts outstanding (in millions), experienced an exponential growth in their time period of analysis January 1998–June 2006. Domanski and Heath point to the increasing presence of financial investors in commodity markets and how this trend has contributed to an increase in commodity market liquidity. With the increase in commodity prices during the study period, they observe a greater variety of financial investors and investment strategies in commodity markets, with passively managed investment and portfolio products being one of the areas of rapid growth. They document that by mid-2006, about \$85 billion of funds were tracking the GSCI and the Dow Jones/AIG Index (two of the most popular commodity indexes financial investors follow). Similarly, the presence of hedge funds, which have a short-term focus, increased; for example, funds that were active in energy markets had tripled to more than 500 since the end of 2004, with an estimated \$60 billion in assets under management. Other examples of financial investor interest in commodity markets included

the number of exchange-traded funds (ETF), the introduction of complex trading strategies (e.g., cross-market arbitrage), managed money traders (MMTs) in oil and natural gas, and the increased volume of OTC transactions.

The first econometric investigation of the relationship between stocks and commodities, using a relative strength measure as in Banister and Forward, is found in Zapata et al. based on annual values of the S&P 500 and the U.S. producer price index (PPI) for the period 1871–2010. They provide a detailed analysis of the history of commodity markets and their relationship to economic growth in the U.S., farm policy, the value of the dollar, inflation, energy markets, farm credit, and bio-energy policy. As in this study, Zapata et al. used the CF bandpass filter by setting minimum and maximum values of periodicity consistent with observations made in previous studies (e.g., Banister and Forward) for commodities (18 and 36, respectively), and minimum and maximum values typical of business cycles (2 and 8 years). The analysis was carried out on the relationship between stocks and all commodities, stocks and farm products, stocks and food products, and stocks and metals. Defining the length of a cycle as the time it takes to move from peak to peak, it was found that stocks and commodities alternated in price leadership with cycles of length 29–32 years, an average of about 31 years, and this empirical regularity was similar in shape across the various commodity groups but somewhat changed in frequency and amplitude across them, thus lending support to the heterogeneity in commodities claimed by previous studies. Zapata et al. also estimated a risk-return model of commodities and individual stocks and found that, in consistency with the stock-commodity cycle, there are periods when investment interest moves with the cycles identified in their paper. While Zapata et al. analyzed various commodity groups (farm, food, energy, and metal markets), they did not study whether the estimated cyclical behavior was applicable to tradable commodity indexes. Nonetheless, the study of the stocks-commodity cyclical phenomenon became of interest to many other researchers (e.g., Chen et al. 2015; Wang et al. 2017; Galariotis and Karagiannis 2021; Hernandez et al. 2021; Reboredo et al. 2021; among many others).

3. Materials and Methods

A relative price strength measure (RS) is used to compare the long-run cyclical relationship between stock and agricultural-commodity prices (Bannister and Forward 2002). The RS compares the relative price performance of stocks and agricultural-commodity prices. The RS is computed by dividing a stock market index into a commodity index for the period 1871–2022. Banister and Forward proposed the use of the S&P 500 index to represent the broad market index in the U.S. and the Producer Price Index in the U.S. (PPI) to represent commodity market prices. When the RS moves up, it indicates that stock returns are outperforming commodity returns (a bull trend in stocks). When the RS is trending down, RS points to commodity returns outperforming stock returns (a bull trend in commodities). In this application, and consistent with commodity investment practice, the Bloomberg Commodity Index (BCOM) and the S&P Goldman Sachs Commodity Index (SPGSCI) are also used to measure the relationship to broad financial markets of two tradeable commodity indexes. While many other tradeable commodity indexes are presently used in commodity investing, we focus on the BCOM and SPGSCI since these indexes have annual data of sufficient length for comparison to their aggregate counterparts, which according to the literature, may have cycles that can last for several decades.

3.1. Measuring Cycles

The proposition of this paper is that certain frequency components of the pricing relationship between stock and commodity markets can be key inputs of investment policy. In this context, it is best to refer to it as a stock-commodities cycle phenomenon and define it as the recurrent fluctuations in the relationship between the two asset classes.² The Christiano and Fitzgerald (2003) bandpass filter, referred to as CF hereafter, is used in this research due to its optimality in modeling time series data, which may or may not be stationary. The CF approximates an ideal bandpass filter for a time series x_t and computes

its cyclical and trend components. A comprehensive treatment of cycles is found in a special issue of the *Journal of Applied Econometrics* and particularly in the article by Harding and Pagan (2005). Briefly, the CF methodology considers a stochastic process x_t , $t = 1, 2, 3, \dots, T$, which has an orthogonal decomposition given by:

$$x_t = y_t + \tilde{x}_t \tag{1}$$

The process y_t has power only in frequencies belonging to an interval I_1 in $(-\pi, \pi)$ and \tilde{x}_t has power only in the complement of I_1 in $(-\pi, \pi)$ with subinterval constants a and b bounded as $0 < a \leq b \leq \pi$; defined this way, $I_1 = \{(a, b) \cup (-b, -a)\}$ with $a = 2\pi/p_u$ and $b = 2\pi/p_l$. The period of oscillations (periodicities of y_t) falls between p_l and p_u , with $2 \leq p_l < p_u < \infty$; thus, in this decomposition, the time series can be written as in Equation (1). The filter weights are chosen to minimize a mean squared error function of y_t and the filtered y_t . The CF filter is a finite sample approximation to the ideal bandpass filter and is given by

$$y_t = \hat{B}(L)x_t \tag{2}$$

where the weights $\hat{B}_{t,j}$ of the approximation are a solution to

$$\hat{B}_{t,j} = \arg \min E \left\{ (y_t - \hat{y}_t)^2 \right\}. \tag{3}$$

In this definition the CF filter is a finite data approximation to the ideal bandpass filter and minimizes the MSE in Equation (3), accommodates unit-root processes for non-stationary time series, and provides a good approximation in the case of stationary time series.³ Christiano and Fitzgerald provide a comparison of the performance of the CF filter relative to other popular filters such as the Baxter-King filter and the Hodrick-Prescott filter;⁴ they also make the filter available in various econometric packages (e.g., EViews, STATA, RATS, R). The bounds for the CF filter in this article were set to coincide with the empirical regularity found in the literature (Bannister and Forward 2002) that stocks and commodities alternate in price leadership for an average of 18 years. Following standard practice, the estimated length of a cycle is defined as the time it takes from peak-to-peak or through-to-through, and since RS is a measure of relative pricing performance, the increasing phase of cycles is a measure of the length of time (in years) when stocks perform better than commodities and vice versa.

3.2. The Data

Annual average data of the Standard and 'Poor's (S&P 500) from 1871 to 2022 was calculated from the monthly observations from Shiller (2022), and the corresponding data needed for the annual PPI for all commodities was obtained from the Federal Reserve Bank of Saint Louis (FRED) and the Bureau of Labor Statistics (U.S. Bureau of Labor Statistics 2022). The impact of the 2008 financial crisis and the financial stress caused by the COVID-19 pandemic is also analyzed and changes in the relationship between stocks and commodity markets are summarized; monthly values of the PPI and S&P 500 from July/2007–June/2010 and January/2019–December/2021 were used for this purpose. The base year for the PPI used in this study is 1982 = 100.⁵

In addition to the PPI, the S&P 500 was compared to two other commodity indices: the Bloomberg Commodity Index (BCOMa futures-based index) from 1960–2022 and the S&P GSCI (SPGSCI a commodity index used as a benchmark for commodity investments) from 1970–2022. Data for both the BCOM and the SPGSCI were downloaded from Bloomberg. The S&P 500 was chosen as a benchmark for the aggregate stock market, which encompasses about 80% of available market capitalization (S&P Dow Jones Indices 2022). Note that except for Cocoa, all of the commodities included in the S&P GSCI index are also included in the BCOM. However, the BCOM includes many other commodities such as livestock, minerals, natural gas, and oil products.

4. Results

4.1. Impact of Financial and COVID-19 Crises

The effect of the 2008 financial crisis (left chart) and the COVID-19 pandemic (right chart) on the RS performance of stock and commodity markets is shown in Figure 1, using monthly values of the PPI and S&P 500 from July/2007–June/2010, and January/2019–December/2021. The financial crisis of 2008 started when foreclosure rates doubled in December 2007. While a number of macro-measures were adopted to stop the roller coaster that was about to unfold, one of the first decisions was U.S. President Bush’s signing of the Tax Rebate Act in February 2008. As the crisis unfolded with bank failures, bailouts by the Federal Reserve started in March 2008, and, of course, the bankruptcy of the Lehman Brothers created global panic, leading to the lowest point in the RS, the time at which the U.S. Dow Jones Industrial Average had a total decline of 53.4% (Figure 1, left chart). Naturally, all of these events altered the RS’s short-term trajectory, marking the end of the bull market in commodities and marking the start of a new period of growth in stock markets. The recovery phase from the financial crisis was much slower and less steep than what occurred during the COVID-19 Pandemic (Figure 1, right chart). In fact, both crises may be characterized by a “V” shape pattern, the major difference being the speed of recovery (i.e., recovery was faster during the Pandemic).

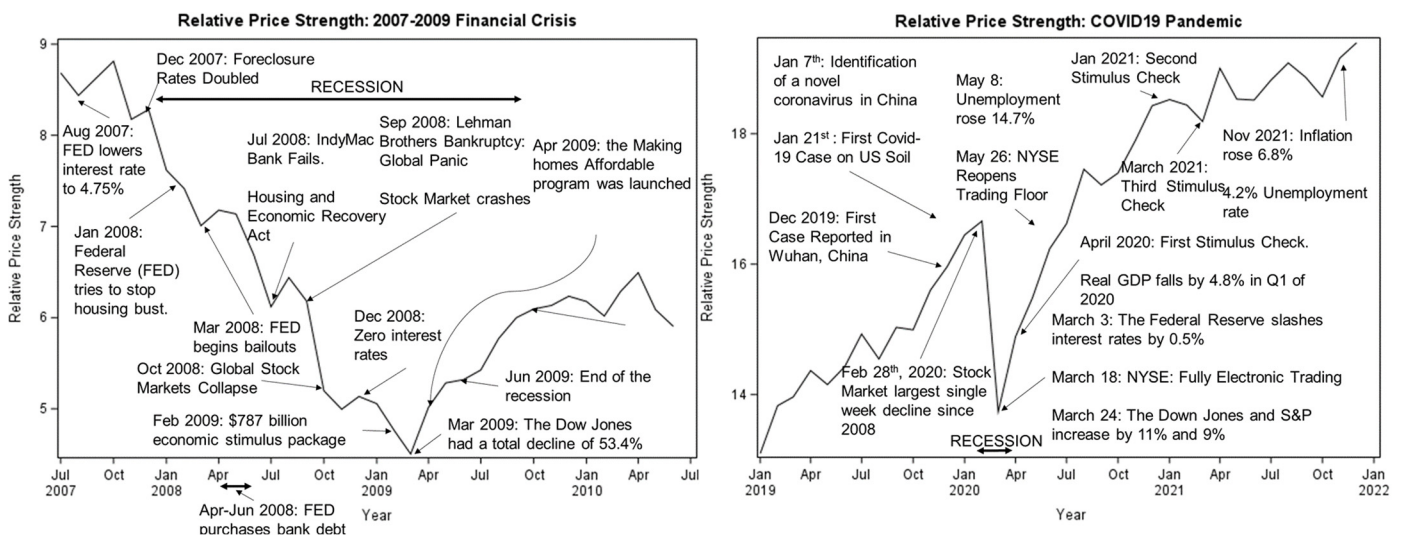


Figure 1. Stock market (S&P 500) versus PPI for all commodities during the 2008 financial crisis and the COVID-19 pandemic.

Although the effects of the crises on unemployment and Gross Domestic Product were more severe during 2020 than during the 2008 financial crisis (Verick et al. 2022), prompt health actions and fiscal policies helped the recovery from the COVID-19 crisis at a much faster pace than that of the 2008 financial crisis (note the differences in the “V shapes”). The overriding conclusion is that the COVID-19 pandemic did not change the long-term trend of the RS between stock and commodity markets and was less impacting than the 2008 financial crisis. The extent to which the pandemic disrupted the workings of global economies, particularly the impact on supply chains, is an issue of continued analysis as of the writing of this paper. Unquestionably, however, the impact on the stability of global financial markets has been severe (e.g., Tan et al. 2022). The impacts of the coronavirus pandemic have depressed market returns for certain U.S. crop farmers and the impacts have varied depending on specific geographic areas and have worsened with the continued disputes with China. Nonetheless, ad hoc U.S. federal aid has helped support farm incomes in areas such as the U.S. Midwestern grain farms (e.g., Schnitkey et al. 2021).

4.2. Commodity Cycles

Figures 2–4 show the results of the Christiano-Fitzgerald bandpass filter applied to the relative price strength (RS) of (a) stocks and the producer price index for all commodities (Figure 2), (b) stocks and the Bloomberg Commodity Index (Figure 3), and (c) and stocks and Goldman Sachs Commodity Index (Figure 4). Figure 2 shows the RS cycle over the period 1871–2022. Note that variability in the RS cycle began to increase approximately in the 1930s and continues to the present. The overall upward tendency of the RS cycle observed in Figure 2, Panel A confirms that on average, stock returns outperformed commodities over the last 151 years. However, the Christiano-Fitzgerald econometric cycle (Figure 2, Panel B) demonstrates the cyclical behavior of the relative pricing performance of stock-commodity markets. Based on Figure 2, Panel B, three cycles (from peak to peak) can be fully identified. The first cycle lasted 24 years, and the subsequent cycles had a length of 35 and 34 years, respectively. The average length of the three stock commodity cycles shown in Figure 2 is 31 years. Also, it should be noted that there is a fourth cycle in progress that began 23 years ago, whose peak is yet to occur but is quickly approaching it.

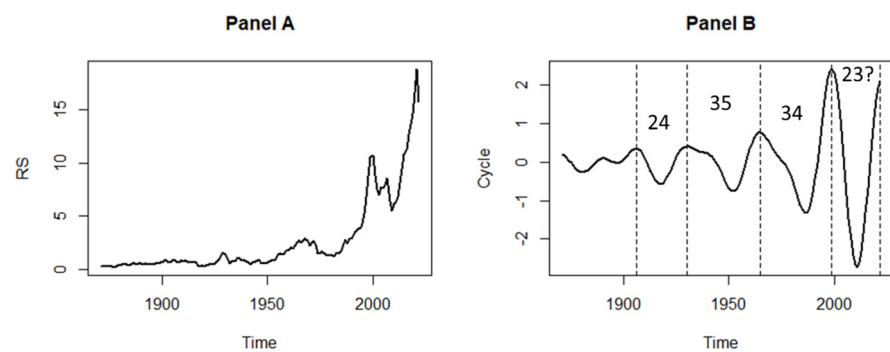


Figure 2. Stock market (S&P 500) versus PPI for all commodities Relative Strength (Panel A) and the RS CF cycle (Panel B), Annual 1871 to 2022.

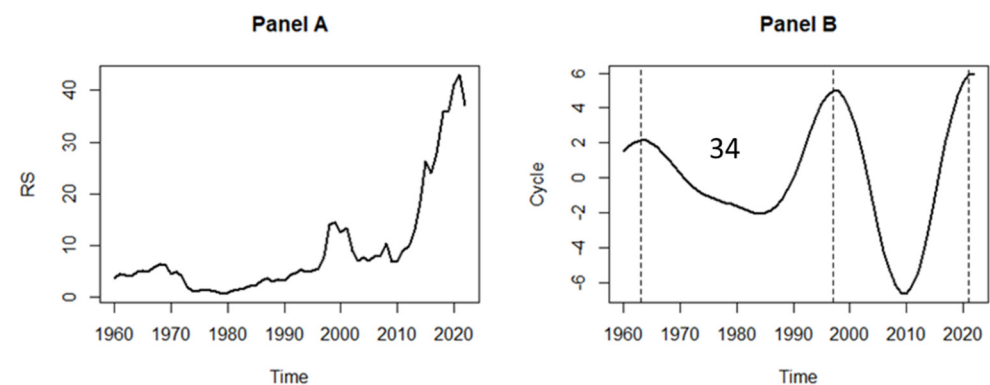


Figure 3. Stock market (S&P 500) versus BCOM Relative Strength (Panel A) and the RS CF cycle (Panel B), Annual 1960 to 2022.

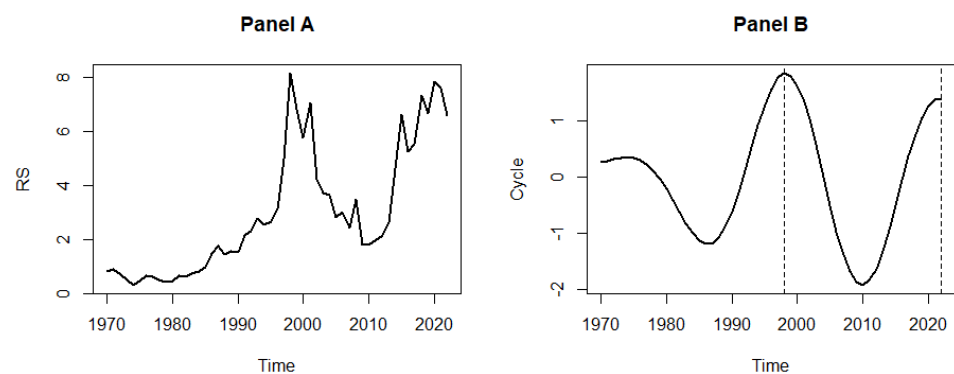


Figure 4. Stock markets (S&P 500) versus the S&PGSCI Relative Strength (Panel A) and the RS CF cycle (Panel B), Annual 1970 to 2022.

The cyclical pattern in RS shown in Figure 2 has important implications for investors interested in commodity investing. Summary statistics tabulated from the indexes used in this article show that during the previous 151 years, the average annual returns of the S&P 500 and the PPI were 4.5% and 1.9%, respectively. But when such measures were generated by segmenting the data according to the RS cycles, it was found that when stocks outperformed commodities (uptrend phases of the RS cycle), stock annual returns were 9.25% while commodities annual returns were 0.15%. Inversely, when commodities outperform stocks (downtrend phase of the RS cycle), stock returns were -1.6% , and commodity returns were 4.9% annually. It is evident from Figure 2 that the benefit of commodity diversification may lie in their “cyclical” nature. The abundant literature promoting commodity investing generally concludes with phrases such as adding commodities provides diversification benefits during “contractions in the stock market”, which is consistent with what the RS cycle of Figure 2 reveals.

To provide a more investing-related comparison between stocks and commodities, we compared the linkage between the S&P 500 and the two investable commodity indexes, the BCOM and the S&P GSCI. Figure 3 displays the relative performance between the S&P 500 and the Bloomberg Commodity Index. The results in Figure 3, Panel B confirm that the mid-1960s to the early 1980s and during most of the decade from 2000 to 2010, were periods of high growth in commodity prices. This is consistent with historical data which show that commodity prices experienced a significant surge during the 1970s due to geopolitical events such as the oil embargo, and again in the early 2000s due to the strong demand from emerging markets such as China and India. The length of the cycle shown in Figure 3, Panel B is 34 years, which is equal to the last full cycle in Figure 2, Panel B. However, the amplitude of the cycle gets larger when BCOM is used as a proxy for commodities and is also pointing to an end in the RS expansion for stocks.

Figure 4 (Panel A) shows the price performance of the S&P 500 relative to the S&P GSCI. Launched on 11 April 1991, the S&P GSCI is a well-known benchmark for representing global commodity prices. Even though the annual index data available starting in 1970, it can be seen in Panel B that the results are very similar when using the S&P GSCI and the BCOM, which is natural given that both indices track many of the same commodities (but not the same markets).

Overall, Figures 2–4 support previous literature concerning the countercyclical behavior of commodities and the stock markets. The returns in the stock market exceed the returns in the commodity markets on average. However, when investment horizons are set to the length of the RS cycle, returns from investing in indexes, such as the Bloomberg Commodity Index and the S&P Goldman Sachs Commodity Index, are higher.

5. Discussion

Applying the bandpass filter proposed by Christiano and Fitzgerald to a measure of pricing performance (RS), it is found that commodity markets, either at the aggregate level

(PPI) or at the more disaggregate level (tradeable commodity indexes), follow a strong cyclical pattern with the broad market, measured by the S&P 500, that lasts about 31 years from peak to peak. This finding is consistent with the findings in previous research (e.g., Zapata et al. 2012; Bannister and Forward 2002). Naturally, the results point to alternating pricing-performance leadership over which financial markets dominate in pricing performance relative to commodities over long periods of time, and as found in other research, the length of the alternating leadership in the cycle (from peak to trough) can run for around 15 years. More specifically, the uptrend phase of the RS cycle indicates that stock returns are outperforming commodity returns. Similarly, the downtrend phase of the cycle indicates that commodity returns outperform stock returns. Data from the past 151 years provide strong evidence of four occasions on which, on average, commodities have outperformed stocks: 1907–1920, 1930–1938, 1969–1982, and 2000–2009. In essence, commodity booms happen when an unanticipated shock increases the demand for certain commodities, while the supply of those commodities takes time to determine prices. Eventually, as supply increases in response to higher prices, the cycle goes into a downturn again, which is often referred to as a “bust” (Büyüksahin et al. 2016). The rise in the price of commodities is associated with wars, inflationary periods, oil prices, and other factors that in one way or another favor the increase in prices. Certainly, the increase in the price of commodities can provide benefits for producers (e.g., farmers and metal producers) and exporting countries. However, it should not be ignored that it can also have devastating effects on those countries that depend on commodity imports and on the purchasing power of middle- or low-income consumers. From a commodity investing perspective, the cyclical interplay between stocks, as measured by the S&P 500, versus tradeable commodity indexes, merits portfolio consideration.

6. Conclusions

Since 2009, stock returns have dominated returns of commodities. However, should the cyclical phenomenon reported here continue, commodity returns will likely outperform stock returns and will continue to attract investors in the coming years. As discussed in the review of literature, wars tend to precede commodity booms. The ongoing Russian-Ukrainian war could lead to a long period of rising commodity prices. Both countries are very important players in the energy commodities, metals, and grain markets. Even if this war does not escalate to a major world conflict, its end is still uncertain, therefore a precautionary buildup of commodity inventories could trigger the next commodity price boom. Another reason for the next commodity boom could be the increasing food demand caused by a growing population. According to some estimates, the world population reached 8 billion in November 2022. This growth in the world population directly increases the demand for food commodities. A strong China recovery would also have a similar impact on commodity prices.

Whether commodity markets are “the world’s best market” (Rogers) or a market of “disappointing return” (Irwin et al. 2020) is a debate that will continue to exist. However, this article provides strong evidence that the real benefits of investing in commodities may lie in their cyclical behavior relative to stocks. Not only do commodities move over time in ups and downs in response to market fundamentals, but their cyclical behavior also tends to co-move opposite to the cycle in stocks. Therefore, an investor who follows the RS cycle can choose commodities for diversification as a hedge. To do so effectively, risk and the phase of the cycle must be accounted for. Similarly, the covariance between the two asset classes must be dynamically integrated with the cyclical relationship, a topic which is currently under investigation.

One area for future research⁶ on this subject is the application of econometric forecasting models to predict turning points in the RS between broad financial markets and commodity investments. Recent developments in the application of machine learning models to prediction problems in econometrics may offer paradigms for more accurate measurement of the numerous factors that impact the risk-reward relationships between

stock and commodity markets and the prediction of turning points (e.g., Giusto and Piger 2017). It should also be noted that the cyclical results in this article suggest commodity leadership for the near future, and this result is consistent with the current forecast given by some market analysts. For instance, Goldman Sachs projects that the S&P GSCI, its commodities tracking index, would increase by up to 43% in 2023. The scarcity of metals and energy-related commodities would be the reason for this price spike. Stockpiles are depleting, and markets are constrained as a result of a lack of investment in mining and the search for new oil sources. Once the US and China's economies recover from their recent economic turbulence, Goldman Sachs anticipates that the commodity market will see a boost in prices (Wallace 2022).

Thinking of commodities as heterogeneous in boom-and-bust periods implies that adding commodities to traditional portfolios of stocks and bonds must account for the difference in commodity responses to increased demand, prices, and investments. Goldman Sachs (Wallace 2022) states that underinvestment in commodity markets precedes a bullish sentiment in commodities, and that despite broadly depleted working inventories and sparse capacity nearly exhausted across most markets in commodities such as oil, capital in 2022 was not responsive to near record prices as market positioning leaned towards a recession (a point also highlighted in Banister and Forward and Radezki). Underinvestment in commodity markets, a disorderly reopening of the Chinese economy, and the rising cost of capital reduce the likelihood of sequential growth in 2023. If commodity markets remain in a state of long-run shortages with subsequent higher and more volatile prices as Goldman Sachs claims and given the recent pause in Fed rate hikes in the U.S. and the impact from the Chinese economy reopening, the leadership in commodity market returns is likely in the foreseeable future, as suggested by the estimated cycle in this paper, which at the end of 2022 is getting to close to a peak. Heterogeneity in commodity markets also implies that, for commodities such as agriculture, the response to, for instance, investments will naturally be faster than would be the case for oil and metals. In both markets, however, the nature of the response will be dictated by the cost of capital.

While the study of these research results was in progress, Goldman Sachs reaffirmed its prediction that commodity markets will be dominated by underinvestment in early 2023. The high cost of capital caused by the rise in interest rates (a deflationary action) has discouraged investors from holding commodity inventories, which could drive commodity prices higher. The cost of capital has also been linked to the withdrawal of more than \$100 billion from commodity ETFs, active mutual funds, and the Bloomberg Commodity Index. What is even more worrisome according to Goldman Sachs is the underinvestment in production which leads to a reduction of commodity inventories, removing a key buffer against shocks in commodity prices. Underinvestment alone does not generate a price shock in commodity markets. Instead, it increases the sensitivity of the commodity markets to demand shocks. Cycle phenomena tend to be popular with some investors and this paper has found that using an approximation to an optimal bandpass filter, aggregate commodity market indexes (PPI) and for tradeable commodity indexes (SPGSCI and BCOM), the phenomenon repeats on average about every 31 years, in consistency with previous work on the subject (e.g., Zapata et al. 2012; Gorton and Rouwenhorst 2006). The indexes used in this study reflect the heterogeneity in commodities that tends to appeal to investors; nonetheless, the investment merits of commodities require a closer examination (e.g., Kat et al.), and preliminary work on this for global markets is beginning to emerge (e.g., Hernandez et al.). Wallace (2022) argues that an underinvestment cycle is at work in the 2023 commodity outlook. Capital expenditures and their relationship to capital consumption are drivers of the supply response in commodity markets. The relationship between these two factors and how they may contribute to the stock-commodity cyclical phenomenon in investing is an area of future research.

Author Contributions: All authors contributed equally to this work. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported in part by the USDA National Institute of Food and Agriculture, Hatch Project LAB94491, accession number 1023839.

Data Availability Statement: Data are available upon request.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

Notes

- ¹ The U.S. producer price index (PPI) is a measure of the average change over time in selling prices received by domestic producers of goods and services. PPIs are available for all commodities and for various producing sectors of the U.S. economy (e.g., mining, manufacturing, agriculture, fishing, forestry) as well as natural gas, electricity, construction, etc. (<https://fred.stlouisfed.org/series/PPIACO> accessed on 20 November 2022). The S&P GSCI index is a benchmark for investment in the commodity markets and is a tradeable index available to market participants of the CME (it is the first major investible commodity index; the Bloomberg ticker symbol is SPGCC1—<https://www.spglobal.com/spdji/en/indices/commodities/sp-gsci/#overview> accessed on 20 November 2022). The Bloomberg Commodity Index (ticker symbol BCOM—<https://www.bloomberg.com/> accessed on 20 November 2022) provides a broad-based exposure to physical commodities via commodity futures contracts, and no single commodity or commodity sector dominates the index. Components of all the above indexes can be found at the links provided above.
- ² This view is consistent with the theory of business cycles (e.g., Prescott 1986; Lucas 1977). If we knew exactly which theoretical model best represents the economy of the U.S., then we could derive theoretical findings for cyclical behavior. As discussed in Prescott, if markets did not display this cyclical phenomenon, it would be puzzling no matter what the true model of the economy is.
- ³ Test of unit-roots, using the augmented Dickey-Fuller statistics with a drift, indicated that RSt has a unit-root, and therefore, the CF filter was specified with a drift and a unit-root. Cointegration tests between the S&P500 and the PPI, resulted in no cointegration for the study period 1960–2022.
- ⁴ Wavelet analysis is an alternative approach to cycle analysis. Cyclical behavior and its components can be analyzed where higher levels of precision of the wavelet approximations are constructed so that it effectively produces a band pass filter that isolates the cycle. The Christiano-Fitzgerald bandpass filter used here to isolate the cycle, which covers both stationary and nonstationary processes, is a similar procedure (e.g., Bowden and Zhu 2008).
- ⁵ We use the base 1982 = 100 because this is the measure that is most widely available in the U.S. (see <https://fred.stlouisfed.org/> accessed on 20 November 2022 and search for PPI). Most research in the U.S. use this base in their work. See Tomek and Robinson (1991) for details.
- ⁶ Another potential topic of future research is the application of theories that fit long-term cycles such as Long Wave Theories. We thank an anonymous reviewer for pointing this out.

References

- Bannister, Barry, and Paul Forward. 2002. The Inflation Cycle of 2002 to 2015. *Legg Mason Equity Research*, April 19.
- Beenen, Jelle. 2005. Commodities as a strategic investment for PGGM. In *An Investors Guide to Commodities*. London: Deutsche Bank AG London, Winchester House.
- Bhardwaj, Geetesh, and Adam Dunsby. 2013. The Business Cycle and the Correlation between Stocks and Commodities. *The Journal of Investment Consulting* 14: 14–25. [CrossRef]
- Bhardwaj, Geetesh, Gary Gorton, and Geert Rouwenhorst. 2015. *Facts and Fantasies about Commodity Futures Ten Years Later*. No. w21243. Cambridge: National Bureau of Economic Research.
- Billah, Mabruk, Faruk Balli, and Indrit Hoxha. 2023. Extreme connectedness of agri-commodities with stock markets and its determinants. *Global Finance Journal* 56: 100824. [CrossRef]
- Bodie, Zvi, and Victor I. Rosansky. 1980. Risk and Return in Commodity Futures. *Financial Analysts Journal* 36: 27–31. [CrossRef]
- Bowden, Roger, and Jennifer Zhu. 2008. The agribusiness cycle and its wavelets. *Empirical Economics* 34: 603–22. [CrossRef]
- Büyükşahin, B., K. Mo, and K. Zmitrowicz. 2016. Commodity Price Supercycles: What are they and What lies ahead? *Bank of Canada Review-Autumn* 35–46.
- Chen, Songjiao, William W. Wilson, Ryan Larsen, and Bruce Dahl. 2015. Investing in agriculture as an asset class. *Agribusiness* 31: 353–71. [CrossRef]
- Christiano, Lawrence J., and Terry J. Fitzgerald. 2003. The Band Pass Filter. *International Economic Review* 44: 435–65. [CrossRef]
- Domanski, Dietrich, and Alexandra Heath. 2007. Financial Investors and Commodity Markets. *BIS Quarterly Review* 53–67.
- Evans, Mark, and Andrew C. Lewis. 2005. Dynamics metals demand model. *Resource Policy* 30: 55–69. [CrossRef]
- Fabozzi, Frank J., Roland Fuss, and Dieter G. Kaiser. 2008. *The Handbook of Commodity Investing*. New York: John Wiley & Sons.
- Galariotis, Emilios, and Konstantinos Karagiannis. 2021. Cultural dimensions, economic policy uncertainty, and momentum investing: International evidence. *The European Journal of Finance* 27: 976–93. [CrossRef]

- Giusto, Andrea, and Jeremy Piger. 2017. Identifying business cycle turning points in real time with vector quantization. *International Journal of Forecasting* 33: 174–84. [CrossRef]
- Gorton, Gary, and K. Geert Rouwenhorst. 2006. Facts and fantasies about commodity futures. *Financial Analysts Journal* 62: 47–68. [CrossRef]
- Greer, Robert. 2000. The Nature of Commodity Index Returns. *The Journal of Alternative Investments* 3: 45–52. [CrossRef]
- Harding, Don, and Adrian Pagan. 2005. A Suggested Framework for Classifying the Models of Cycles Research. *Journal of Applied Econometrics* 20: 151–59. [CrossRef]
- Hernandez, Jose A., Sang H. Kang, and Seong-Min Yoon. 2021. Spillovers and portfolio optimization of agricultural commodity and global equity markets. *Applied Economics* 53: 1326–41. [CrossRef]
- Irwin, Scott H., Dwight R. Sanders, Aaron Smith, and Scott Main. 2020. Returns to Investing in Commodity Futures: Separating the Wheat from the Chaff. *Applied Economic Perspectives and Policy* 42: 583–610. [CrossRef]
- Kat, Harry M., and Roel C. A. Oomen. 2006. What every investor should know about commodities, Part II: Multivariate return analysis. *Alternative Research Centre Working Paper* 33. [CrossRef]
- Lucas, Robert E. 1977. Understanding business cycles. *Journal of Monetary Economics* 3: 7–21. [CrossRef]
- Prescott, Edward C. 1986. Theory ahead of business-cycle measurement. In *Carnegie-Rochester Conference Series on Public Policy*. Amsterdam: North-Holland, vol. 25, pp. 11–44.
- Radetzki, Marian. 2006. The anatomy of three commodity booms. *Resources Policy* 31: 56–64. [CrossRef]
- Reboredo, Juan Carlos, Andrea Ugolini, and Jose Arreola Hernandez. 2021. Dynamic spillovers and network structure among commodity, currency, and stock markets. *Resources Policy* 74: 102266. [CrossRef]
- Rogers, Jim. 2004. *Hot Commodities: How Anyone Can Invest Profitably in the World's Best Market*. New York: John Wiley & Sons.
- Rouwenhorst, K. Geert, and Ke Tang. 2012. Commodity Investing. *Annual Review of Financial Economics* 4: 447–67. [CrossRef]
- S&P Dow Jones Indices. 2022. S&P 500. Available online: <https://www.spglobal.com/spdji/en/indices/equity/sp-500/#overview> (accessed on 20 November 2022).
- Schnitkey, Gary D., Nicholas D. Paulson, Scott H. Irwin, Jonathan Coppess, Bruce J. Sherrick, Krista J. Swanson, Carl R. Zulauf, and Todd Hubbs. 2021. Coronavirus Impacts on Midwestern Row-Crop Agriculture. *Applied Economic Perspectives and Policy* 43: 280–91. [CrossRef] [PubMed]
- Shiller, Robert. 2022. Available online: <http://www.econ.yale.edu/~shiller/data.htm> (accessed on 20 November 2022).
- Tan, Xiaoyu, Shiqun Ma, Xuetong Wang, Chao Feng, and Lijin Xiang. 2022. The impact of the COVID-19 pandemic on the global dynamic spillover of financial market risk. *Frontiers in Public Health* 10: 963620. [CrossRef] [PubMed]
- Tomek, William G., and Kenneth L. Robinson. 1991. *Agricultural Product Prices*. Ithaca: Cornell University Press.
- U.S. Bureau of Labor Statistics. 2022. Available online: stats.bls.gov (accessed on 20 November 2022).
- Verick, Sher, Dorothea Schmidt-klaus, and Sangheon Lee. 2022. Is this time really different? How the impact of the COVID-19 crisis on labour markets contrasts with that of the global financial crisis of 2008–09. *International Labour Review* 161: 125–48. [CrossRef] [PubMed]
- Wallace, Paul. 2022. Goldman Says Commodities Will Gain 43% in 2023 as Supply Shortages Bite. *Bloomberg*. Available online: <https://www.bloomberg.com/news/articles/2022-12-15/goldman-commodity-supercycle> (accessed on 2 February 2023).
- Wang, Luo, Bin Li, Rakesh Gupta, Jen-Je Su, and Benjamin Liu. 2017. Return predictability in Australian Managed Funds. *International Journal of Business Economics* 16: 1–19.
- Zapata, Hector O., Joshua D. Detre, and Tatsuya Hanabuchi. 2012. Historical performance of commodity and stock markets. *Journal of Agricultural and Applied Economics* 44: 339–57. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

Article

Oil Price Uncertainty Shocks and Global Equity Markets: Evidence from a GVAR Model

Afees A. Salisu ^{1,2}, Rangan Gupta ² and Riza Demirel ^{3,*}

¹ Centre for Econometrics & Applied Research, University of Ibadan, Ibadan 200132, Nigeria

² Department of Economics, University of Pretoria, Private Bag X20, Hatfield 0028, South Africa

³ Department of Economics and Finance, Southern Illinois University Edwardsville, Edwardsville, IL 62026-1102, USA

* Correspondence: rdemire@siue.edu

Abstract: This paper examines the propagation of oil price uncertainty shocks to real equity prices using a large-scale Global Vector Autoregressive (GVAR) model of 26 advanced and emerging stock markets. The GVAR framework allows us to capture the transmission of local and global shocks, while simultaneously accounting for individual-country peculiarities. Utilising a recently developed model-free, robust estimate of oil price uncertainty, we document a statistically significant and negative effect of uncertainty shocks emanating from oil prices on the large majority of global stock markets, with the adverse effect of oil price uncertainty shocks found to be stronger for emerging economies as well as net oil-exporting nations. Interestingly, however, global stock markets exhibit a great deal of heterogeneity in their recovery following oil uncertainty shocks as some experience rapid corrections in stock valuations while others suffer from extended slumps. While the results are sensitive to the oil uncertainty measure utilised, they suggest that country diversification in the face of rising oil market uncertainty can still be beneficial for global investors as global stock markets exhibit a rather heterogeneous pattern in their recovery rates against oil market shocks.

Keywords: oil price uncertainty shocks; international equity markets; global vector autoregressive model

JEL Classification: C32; G15



Citation: Salisu, Afees A., Rangan Gupta, and Riza Demirel. 2022. Oil Price Uncertainty Shocks and Global Equity Markets: Evidence from a GVAR Model. *Journal of Risk and Financial Management* 15: 355. <https://doi.org/10.3390/jrfm15080355>

Academic Editor: Kentaro Iwatsubo

Received: 27 May 2022

Accepted: 4 August 2022

Published: 9 August 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Building on the pioneering works by Bernanke (1983) and Pindyck (1991), the oil-stock market literature argues that uncertainty and the real options effect associated with investment decisions driven by high oil price uncertainty create cyclical fluctuations in investments by lowering firms' incentives for immediate investment activities. This, in turn, affects the cash flows generated by firms as well as the discount rates used in firm valuations, thus, opening a channel in which oil market uncertainty impacts stock prices and/or returns (Swaray and Salisu 2018; Chen and Demirel 2022). Given the importance of oil as a major input factor for business operations that drive the real economy, its volatility may influence both investment and policy decisions. Thus, uncertainty regarding business profitability, valuations, and investment decisions are all impacted by oil price volatility, which is a source of uncertainty that affects the cost of this essential input (Henriques and Sadorsky 2011).

Shocks emanating from the oil market could have serious implications for the real economy as oil-price-induced uncertainty could force firms to postpone 'irreversible' investment decisions (Elder and Serletis 2009). In other words, uncertainty about the return to investment at the level of the firm may cause cyclical fluctuations in aggregate investments (Bernanke 1983). The foregoing suggests that even when an oil price shock is favourable, the uncertainty about its nature (whether permanent or transitory) may discourage positive investment decisions. Indeed, in related works, Maghyreh and Abdoh (2020) show that oil

market uncertainty has a negative and asymmetric effect on corporate investment, while Yin and Lu (2022) argue that oil uncertainty increases firms' risk-taking through the channel of real options related to firms' growth opportunities. Such a negative effect on investment activities could also be driven by the impact of oil price shocks on the cost of financing firm operations as evidenced by Prodromou and Demirer (2022). Finally, since stock prices can be modelled as the sum of discounted cash flows including dividends, oil price uncertainty can adversely affect stock prices by decreasing the overall profit that a firm generally uses to pay dividends, as firms need to bear additional costs to avoid the risks associated with oil price uncertainty (Demirer et al. 2015).

Accordingly, the literature offers various studies that examine the validity of theoretical arguments that associate oil price uncertainty with stock prices and or/returns via the investment and dividend channels for both the developed (see, Sadorsky (1999); Masih et al. (2011); Alsalman (2016); Diaz et al. (2016); Rahman (2021)) and emerging economies (see, Jiranyakul (2014); Aye (2015); Bass (2017); Benavides et al. (2019)).¹ However, as Chen and Demirer (2022) note, the literature offers mixed evidence at best regarding the effect of oil price fluctuations on stock market dynamics with the majority of the studies documenting a negative oil price effect on stock market returns (e.g., Chen 2009; Basher et al. 2012), while others find insignificant (Demirer et al. 2015; Hatemi et al. 2017) and, in some cases, positive (e.g., Zhu et al. 2011; Silvapulle et al. 2017) oil price effects. Against this background, in contrast with the country-specific/firm-level analyses that are popular in the literature, we contribute to the extant literature by examining the propagation of oil price uncertainty shocks to real equity prices using a large-scale Global Vector Autoregressive (GVAR) model of 26 advanced and emerging stock markets. Among the attractions of the GVAR framework is its ability to accommodate the transmission of local, regional, and global shocks to individual countries, while simultaneously accounting for individual-country peculiarities, thus providing a more comprehensive assessment of the effect of oil market uncertainty on stock market dynamics.

One must realise, however, that uncertainty is a latent variable, and needs to be measured. Given this, the large majority of the above-mentioned studies rely on univariate or bivariate Generalised Autoregressive Conditional Heteroskedasticity (GARCH) models applied to oil price series to derive metrics of oil price uncertainty. In other words, GARCH-based oil price uncertainty is fully determined by changes in the level of oil price, and as a result, it is not possible to disentangle uncertainty about the oil price and changes in the oil price level (Jo 2014). Given this, Rahman (2021) proposes a new measure of oil price uncertainty by utilising Stochastic Volatility (SV) in a Structural Vector Autoregressive (SVAR) model (involving oil and stock prices, and a monetary policy instrument). In this model, oil price uncertainty is the conditional variance of the oil price change forecast error, which evolves independently of any change in the oil price level. Using this framework, Rahman (2021) provides evidence that increased oil price uncertainty has a negative effect on (real) stock returns of the United States.

Despite the innovativeness of the approach adopted by Rahman (2021) over GARCH-based models in measuring oil price uncertainty, the metric is not free from the structure of any specific theoretical model. Given these empirical issues in constructing an appropriate metric of oil price uncertainty, Nguyen et al. (2021) have proposed a novel construction of the oil price uncertainty index that is unconditional on a model. The authors develop a measure of oil price uncertainty as the one-period-ahead forecast error variance of a forecasting regression with SV in the residual terms. The novelty of this approach lies in its flexibility in including a large number of additional information that is important in explaining fluctuations in oil prices namely, exchange rate, oil production, global economic conditions, and co-movement in the fuel market. In this sense, the index is able to capture uncertainty in oil price rather than volatility as measured by both GARCH and SV models. Thus, this feature of the uncertainty metric informs our preference for oil price uncertainty.

Using this more robust metric of oil price uncertainty and building on the mixed evidence in the literature regarding the effect of oil price fluctuations on financial markets,

we aim to extend the literature on the nexus between oil price uncertainty and stock markets by analysing the impact of oil price uncertainty shocks from a global perspective covering 26 developed and developing economies that account for 90% of the global aggregate Gross Domestic Product (GDP). In our empirical application, we rely on the Global Vector Autoregressive (GVAR) framework, originally proposed by Pesaran et al. (2004), which can account for international transmission of shocks (in our case, oil price uncertainty) based on a large panel of country-level macroeconomic data (i.e., output, inflation, short- and long-term interest rates, real exchange rate, over and above the real equity prices) and global exogenous variables (e.g., commodity prices). This framework allows us to analyse the impact of oil price uncertainty on a wide range of global stock markets in a multivariate, simultaneous setting by controlling for a wide array of domestic and global macroeconomic and financial variables, which are known to serve as drivers of international stock market dynamics (Jordan et al. 2016, 2017; Sousa et al. 2016; Aye et al. 2017). In the process, our study provides a more accurate assessment regarding the size of the impact of oil price uncertainty on global equity markets by mitigating the omitted variable bias.

The remainder of the paper is organised as follows: Section 2 presents the data and summary statistics. Section 3 deals with the methodology, while Section 4 is devoted to the discussion of results. Section 5 concludes the paper.

2. Data and Summary Statistics

Given the approach adopted in this study, we consider nine variables in total, out of which six (real gross domestic product, inflation rate, real equity prices, real exchange rate, nominal short- and long-term interest rate) are peculiar to 26 countries covering both developed and emerging economies², while the remaining three (oil price, raw material price, and metal price) represent global/common variables. We use quarterly frequency covering 1979Q1 through 2019Q4, based on the availability of data.³ We present the results of the summary statistics for the domestic (see Table 1) and global variables (see Table 2), while for want of space, only the mean, standard deviation (s.d), and coefficient of variation (cv) statistics are reported.

In terms of real output, India, Singapore, Malaysia, Korea, Chile, Thailand, and the Philippines are found to be the most volatile economies compared to the other countries, judging by the coefficient of variation. On the rate of change in price levels, all economies experience on average, an inflation rate of less than 3 percent with the exception of Argentina, which recorded 14 percent on average during the sample period. For equity prices, South Africa, the United Kingdom, Italy, Sweden, and France recorded the highest equity prices along with Argentina, Finland, Spain, and Korea compared to the rest of the countries. In the commodity market representing all the global variables, oil is found to be the most volatile during the period under study relative to the other commodities.

Table 1. Descriptive statistics of domestic variables.

Country	rgdp			Inf			reqp			rexr			s.int			l.int		
	Mean	s.d	cv	Mean	s.d	cv	Mean	s.d	cv	Mean	s.d	cv	Mean	s.d	cv	Mean	s.d	cv
Argentina	4.64	0.35	0.08	0.14	0.25	1.79	0.23	0.61	2.65	-4.05	0.43	-0.11	0.12	0.19	1.58			
Australia	4.55	0.38	0.08	0.01	0.01	1.00	2.09	0.47	0.22	-4.29	0.40	-0.09	0.02	0.01	0.50	0.02	0.01	0.50
Austria	4.53	0.25	0.06	0.01	0.00	0.00	1.42	0.65	0.46	-4.72	0.37	-0.08	0.01	0.01	1.00	0.01	0.01	1.00
Belgium	4.54	0.22	0.05	0.01	0.01	1.00	1.53	0.56	0.37	-4.74	0.35	-0.07	0.01	0.01	1.00	0.01	0.01	1.00
Canada	4.53	0.30	0.07	0.01	0.01	1.00	1.85	0.59	0.32	-4.32	0.34	-0.08	0.01	0.01	1.00	0.02	0.01	0.50
Chile	4.50	0.57	0.13	0.02	0.02	1.00	2.08	1.29	0.62	1.68	0.31	0.18	0.03	0.03	1.00			
Finland	4.54	0.25	0.06	0.01	0.01	1.00	0.73	0.97	1.33	-4.77	0.29	-0.06	0.01	0.01	1.00			
France	4.53	0.20	0.04	0.01	0.01	1.00	2.31	0.66	0.29	-4.73	0.32	-0.07	0.01	0.01	1.00	0.01	0.01	1.00
Germany	4.54	0.21	0.05	0.01	0.00	0.00	1.46	0.53	0.36	-4.71	0.34	-0.07	0.01	0.01	1.00	0.01	0.01	1.00
India	4.58	0.78	0.17	0.02	0.01	0.50	1.55	0.85	0.55	-1.01	0.29	-0.29	0.02	0.00	0.00			
Italy	4.51	0.15	0.03	0.01	0.01	1.00	2.59	0.44	0.17	-4.74	0.35	-0.07	0.02	0.01	0.50	0.02	0.01	0.50
Japan	4.55	0.19	0.04	0.00	0.01	1.00	1.31	0.37	0.28	0.30	0.38	1.27	0.01	0.01	1.00	0.01	0.01	1.00
Korea	4.42	0.64	0.14	0.01	0.01	1.00	0.63	0.60	0.95	2.39	0.31	0.13	0.02	0.01	0.50	0.02	0.01	0.50
Malaysia	4.46	0.66	0.15	0.01	0.01	1.00	1.57	0.44	0.28	-3.41	0.18	-0.05	0.01	0.00	0.00			
The Netherlands	4.49	0.24	0.05	0.01	0.00	0.00	1.77	0.67	0.38	-4.73	0.35	-0.07	0.01	0.01	1.00	0.01	0.01	1.00
Norway	4.43	0.30	0.07	0.01	0.01	1.00	1.84	0.74	0.40	-2.58	0.35	-0.14	0.02	0.01	0.50	0.02	0.01	0.50
New Zealand	4.87	0.27	0.06	0.01	0.01	1.00	0.45	0.27	0.60	-4.06	0.46	-0.11	0.02	0.01	0.50	0.02	0.01	0.50
Philippines	4.68	0.46	0.10	0.02	0.02	1.00	1.48	1.13	0.76	-0.92	0.34	-0.37	0.02	0.02	1.00			
South Africa	4.65	0.29	0.06	0.02	0.01	0.50	3.90	0.67	0.17	-2.82	0.24	-0.09	0.02	0.01	0.50	0.03	0.01	0.33
Singapore	4.41	0.71	0.16	0.00	0.01	1.00	0.91	0.35	0.38	-4.09	0.38	-0.09	0.01	0.01	1.00			
Spain	4.52	0.28	0.06	0.01	0.01	1.00	0.67	0.68	1.01	-4.76	0.39	-0.08	0.02	0.01	0.50	0.02	0.01	0.50
Sweden	4.57	0.26	0.06	0.01	0.01	1.00	2.37	0.93	0.39	-2.52	0.28	-0.11	0.01	0.01	1.00	0.02	0.01	0.50
Switzerland	4.59	0.20	0.04	0.00	0.01	1.00	1.45	0.77	0.53	-4.22	0.43	-0.10	0.01	0.01	1.00	0.01	0.00	0.00
Thailand	4.51	0.58	0.13	0.01	0.01	1.00	1.47	0.72	0.49	-1.07	0.28	-0.26	0.02	0.01	0.50			
United Kingdom	4.54	0.25	0.06	0.01	0.01	1.00	3.16	0.44	0.14	-4.98	0.36	-0.07	0.01	0.01	1.00	0.02	0.01	0.50
USA	4.50	0.31	0.07	0.01	0.01	1.00	1.82	0.69	0.38	#DIV/0!	0.01	0.01	0.01	0.01	1.00	0.01	0.01	1.00

Note: mean, s.d and cv indicate mean, standard deviation and coefficient of variation (computed as the ratio of s.d and mean) respectively. For domestic variables, rgdp, inf, reqp and rexr are for real gross domestic product, inflation rate, real equity price and real exchange rate (where US dollar is the reference currency), respectively. Furthermore, s.int means nominal short-term rate, while l.int is nominal long term rate.

Table 2. Descriptive statistics of global variables.

Variables	Mean	med	s.d	cv
OIL PRICE	0.69	0.65	1.18	1.710
AGRICULTURAL COMMODITY PRICE	4.58	4.61	5.10	1.114
METAL PRICE	4.46	4.29	5.52	1.238

Note: mean, s.d and cv indicate mean, standard deviation and coefficient of variation (computed as the ratio of s.d and mean) respectively.

3. Methodology

We specify the nexus between oil uncertainty and real stock prices within the GVAR framework of Chudik and Pesaran (2016) that accommodates the transmission of global shocks, such as those associated with oil price uncertainty, to domestic variables, while also accounting for the characteristics of individual economies comprising of both the developed and emerging countries. Thus, the GVAR framework enables us to capture the interconnectedness among the various markets while tracing the propagation of global shocks, which is the main attraction of this modelling framework compared to other multivariate models. The GVAR dataset utilised in our analysis includes quarterly macroeconomic variables for the 26 developed and emerging economies. The sample period is governed by data availability in the updated GVAR dataset maintained by Mohaddes and Raissi (2020) covering 1979Q2 to 2019Q4.⁴

In setting up the GVAR model for the 26 countries, we consider six domestic (endogenous) variables namely, log real GDP, the rate of inflation, short-term interest rate, long-term interest rate, the log real exchange rate, and log real equity prices, and three external (common) factors involving the base metals prices, agricultural commodity prices, and oil price uncertainty.⁵ We, however, focus on the results that highlight the impact of the oil price uncertainty shock on real stock prices by constructing the GVAR model as:

$$x_{it} = \sum_{\ell=1}^{p_i} \eta_{i\ell} x_{i,t-\ell} + \tau_{i0} x_{it}^* + \sum_{\ell=1}^{q_i} \tau_{i\ell} x_{i,t-\ell}^* + \sigma_{i0} \omega_{it} + \sum_{\ell=1}^{r_i} \sigma_{i\ell} \omega_{i,t-\ell} + \mu_{it} \quad (1)$$

where x_{it} is a $k_i \times 1$ vector of variables specific to cross-section unit i ($i = 1, 2, \dots, N$) in period t ($t = 1, 2, \dots, T$); x_{it}^* is the corresponding $k_i^* \times 1$ vector of foreign variables constructed as $x_{it}^* = \sum_{j=1}^N w_{ij} x_{jt}$ where $\sum_{j=1}^N w_{ij} = 1$, and $w_{ii} = 0$; $\eta_{i\ell}$ is a $k_i \times k_i$ matrix of unknown parameters for domestic variables where $\ell = 1, 2, \dots, p_i$; $\tau_{i\ell}$ for $\ell = 0, 1, 2, \dots, q_i$, is a $k_i^* \times k_i^*$ matrix of unknown parameters for foreign variables; $\sigma_{i\ell}$ is a $k_i \times k_i$ matrix of unknown parameters for external common factors for all the cross-sections as $\ell = 0, 1, 2, \dots, r_i$, and μ_{it} is a $k_i \times 1$ vector of error terms. Finally, both the foreign and global factors are treated as weakly exogenous. The GVAR approach to estimating the transmission of shocks involves two steps. First is the estimation of country models as formulated in Equation (1), thereafter, the estimated country models are stacked together to form a large GVAR model from which the impact of oil uncertainty on real stock prices is measured.

4. Empirical Results

4.1. Main Findings

In this section, we present the results of country-specific impulse response functions to assess the impact of shocks emanating from crude oil price uncertainty on real stock prices across the considered 26 economies as captured in our GVAR analysis. The impulse responses of real stock prices to a one standard deviation shock to oil price uncertainty are presented in Figures 1 and 2, respectively, at the individual country and group levels. In each plot, the median response is represented in solid lines while the (16–84%) lower and upper bootstrapped error bands are depicted with dotted lines. In addition, the impacts are measured in percentage by multiplying the values indicated by the solid lines by 100. Considering the recent evidence in Maghyereh and Abdoh (2020) and the

theoretical arguments by Bernanke (1983) and Pindyck (1991), our hypothesis is that higher oil uncertainty should negatively impact stock prices, either from a risk-taking perspective (Yin and Lu 2022) or profitability and cash flow perspectives (Demirer et al. 2015).

We observe in Figure 1 that a one-standard-deviation shock prompted by oil uncertainty causes a statistically significant initial decline in real stock prices in the large majority of the countries in the sample (23 out of the 26). In terms of the magnitude of the impact, consistent with the recent evidence by Chen and Demirer (2022) on global oil exposures, we observe rather heterogeneous effects in the magnitude of the responses, ranging from 1% to 10% with generally larger effects observed for net oil-exporting countries as well as those with high economic complexity as will be discussed in our subsequent analysis. The largest negative impact is observed in the case of S. Korea and Malaysia whereas the effect is found to be insignificantly negative for several countries including Italy, the Philippines, and Spain. Interestingly, however, while the impact of the oil uncertainty shock is generally significantly negative for all economies, the stock market recovery varies across the global markets, suggesting that global stock markets exhibit heterogeneity in how they respond to oil uncertainty shocks in the medium to long runs. Our results, thus, align well with the findings of Chen (2009); Basher et al. (2012); and Alsalman (2016) although none of these works simultaneously assess different levels of economies as in our case. While Chen (2009) and Basher et al. (2012) find that a positive shock to oil prices depresses the emerging market stock prices, Alsalman (2016) reports an insignificant effect of oil price uncertainty on US stock returns.

Most stock markets seem to experience a positive correction in the intermediate term after the initial negative response to an oil uncertainty shock, most notably Chile, for which the market recovers quite steadily over the next several quarters following the shock. In comparison, countries like Argentina, Canada, Malaysia, and Norway seem to experience extended slumps due to the oil market shock, understandably because they have largely oil-dependent economies. The same argument could be made for Japan and especially Korea, for which the oil shock seems to exert extended downward pressure on real stock prices, which can be partly attributed to the drag oil uncertainty exerts on industrial activities with consequences on firm profits and investments. A closer review of some stylised facts suggests that Japan and Korea rank higher in economic complexity⁶ compared to their peers with high oil consumption,⁷ therefore, any shock to oil, which serves as a major input to the unique variety of products, would have far-reaching effects on production (see Salisu et al. 2021) and by extension stock valuations.

Figure 2 presents the impact of a one-standard-deviation shock to oil price uncertainty on various country groups based on developed/emerging and net oil-exporting/importing classifications.⁸ This consideration, particularly for the latter, is motivated by the evidence in the literature suggesting heterogeneous response of stock markets of net oil-exporters and net oil-importers to oil price shocks (see, for example, Wang et al. 2013; Salisu and Isah 2017; among others). The results from country groups in Figure 2 generally confirm the inferences obtained from individual country-level analyses. We find that the shock to oil price uncertainty results in an initial reduction in real stock prices across all country groups, with the impact being felt the most by the net oil-exporting economies. Specifically, a one standard deviation shock to oil price uncertainty results in about 5.4% and 1.7% fall in real stock prices of net oil-exporting and net oil-importing groups, respectively (in the 2nd quarter following the shock), while it results in about 1.8% and 2.3% fall, respectively, for the developed and emerging economies (at the same forecast horizon). While it is not unexpected to find that emerging stock markets have greater exposure to oil price uncertainty and thus respond more to oil uncertainty shocks compared to their developed counterparts, the relatively larger response of net exporters compared to net oil importers is interesting. Clearly, the stock markets whose economies are more dependent on oil are more likely to be negatively impacted by oil price uncertainty than those with less dependence on oil. However, the larger impact on oil exporters could be explained by the lack of diversification in exports and the reliance of those economies on the so-called

petrodollars, which in turn, generates a bigger impact on their stock markets. In comparison, net oil importers exhibit relatively greater resilience to these shocks despite their reliance on oil imports, which can be explained by the relative strength of their import/export diversification. Similarly, emerging markets seem to show some level of resilience against oil uncertainty shocks, as the impact later turns positive, albeit at a long horizon, even though the initial negative impact is stronger than developed equity markets.

From an investment perspective, the results suggest that country diversification in the face of rising oil market uncertainty can still be beneficial for global investors as global stock markets exhibit a rather heterogeneous pattern in their recovery rates to oil market shocks. Therefore, investors who are concerned about rising oil market uncertainty and how this might impact stock market returns can find some comfort in the results such that a dynamic country allocation strategy can not only help to mitigate the negative effects of such shocks but also help improve portfolio returns as these stock markets recover from the shock in the intermediate and long term. For policymakers, our findings can be used as a guide to further examine the role of the economic fundamentals that lead to the observed heterogeneity in the response of these markets to oil price uncertainty shocks.

4.2. Additional Results

4.2.1. The Role of Global Financial Crisis (GFC)

To better appreciate the effect of oil price uncertainty on the real stock prices under different economic conditions, we analyse, in this sub-section, the distinct role of the global financial crisis (GFC) in the oil-stock market nexus. Thus, we partitioned our sample periods into pre- and post-GFC periods (see Figures 3 and 4 for pre-GFC and Figures 5 and 6 for post-GFC). To further allow for more classifications and generalisation of results, we also present the results of the IRFs for different regions including Euro and the G7 countries. Our results show that virtually all the markets seem to be resilient to oil price uncertainty shocks during the pre-GFC as some of them can be used to hedge against the oil price risk. This is observed from the IRFs where they are largely insignificant over a long forecast horizon after a standard deviation shock to oil price uncertainty. This is evident for both the emerging and developed markets including the Euro and G7 countries. However, this pattern observed during the pre-GFC era appears to have largely disappeared after the GFC as global stocks markets respond significantly negatively to the oil price uncertainty shock during this period. Therefore, the negative IRF outcomes reported for the full sample in Figures 1 and 2 are largely driven by the post-GFC period. In other words, the connection between oil price uncertainty and stock markets is episodic and ignoring this feature may lead to biased outcomes.

4.2.2. Robustness Checks with an Alternative Measure for Oil Price Uncertainty

In order to assess the robustness of our findings, we also consider an alternative measure of oil price uncertainty based on the GARCH specification, and the IRFs are presented in Figures 7 and 8 for the individual countries and different groups. Unlike the SV-based oil price uncertainty measure, we find that most of the IRFs are insignificant, and this behaviour is particularly evident for the country groups when the GARCH-based measure is used. As previously espoused in Section 1, the results seem to support the argument raised by Nguyen et al. (2021) that the SV-based oil price uncertainty is a better measure of oil price uncertainty as it appropriately captures the inherent dynamics in the crude oil market. This also highlights the importance of the choice of the metric to capture oil market uncertainty for policymakers as the economic assessment of oil market shocks on the real economy depends on the metric employed in the analysis.

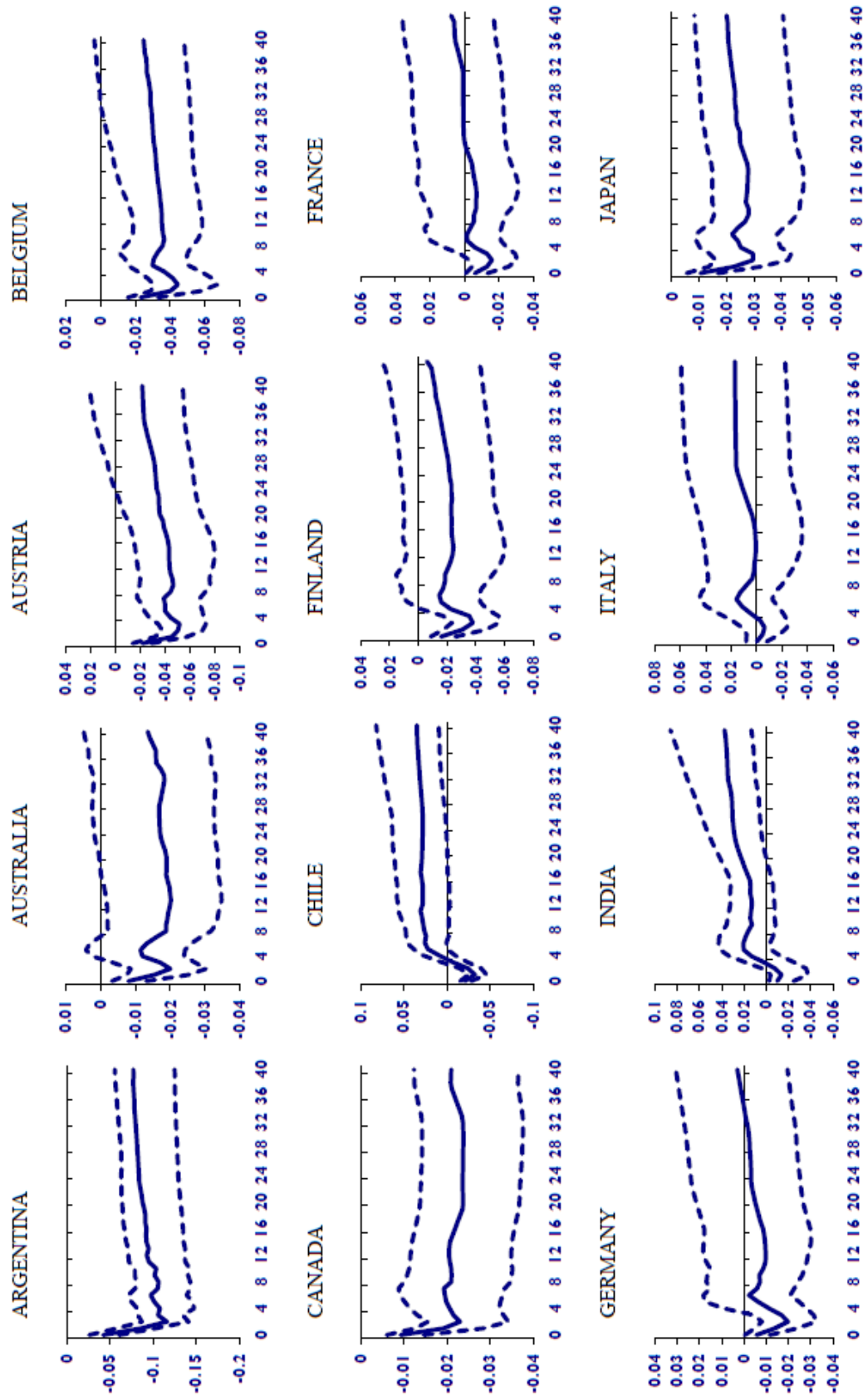


Figure 1. Cont.

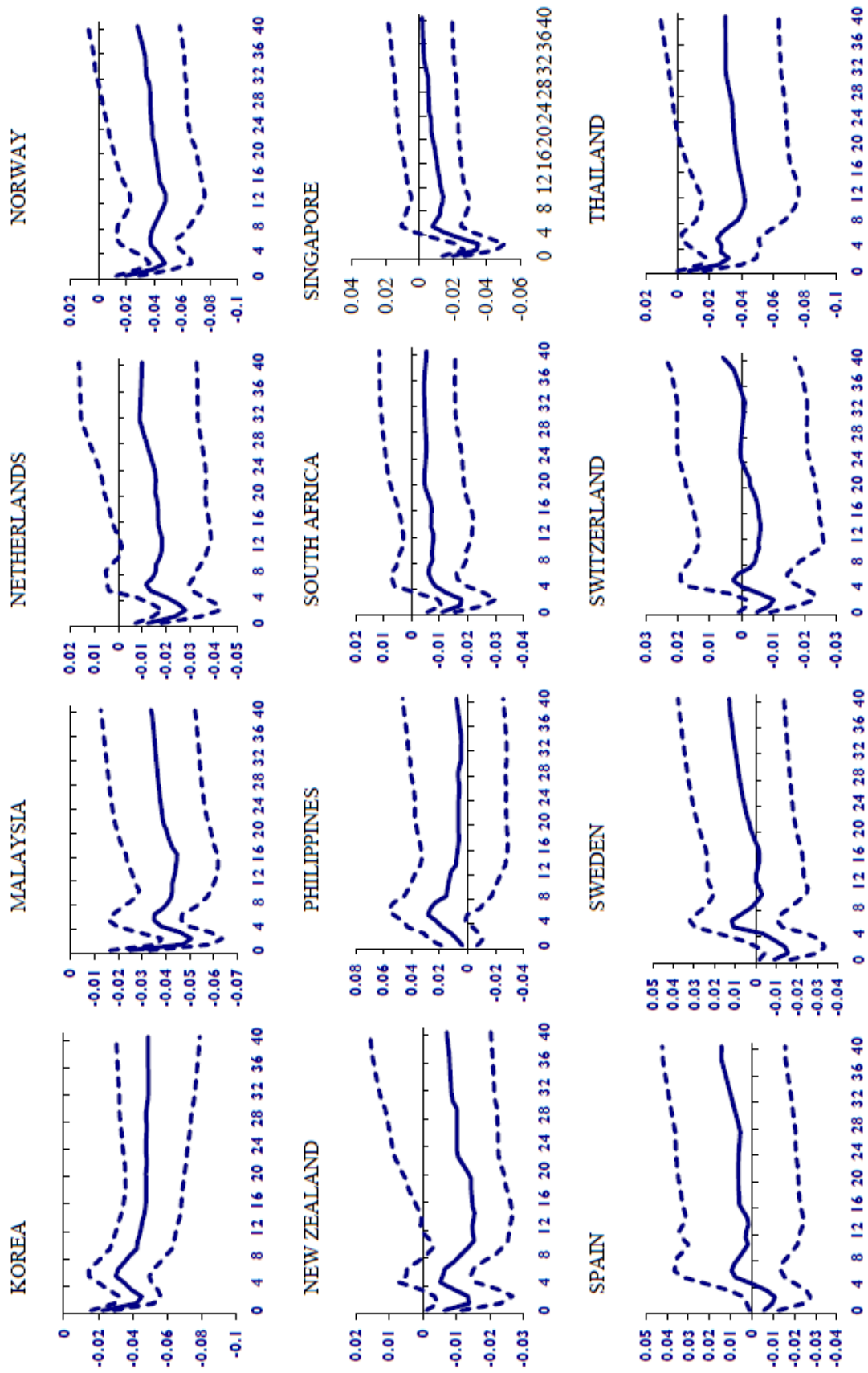


Figure 1. Cont.

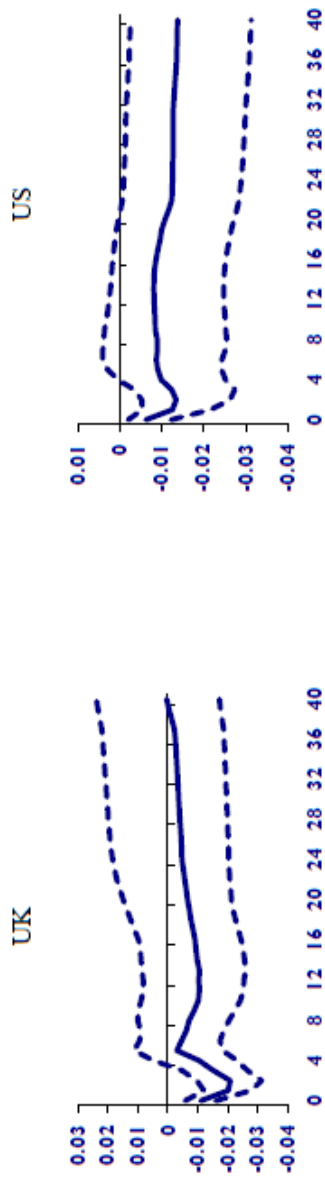


Figure 1. GVAR-based oil price uncertainty shocks and real stock prices (country-level). Note: The figure depicts the country-specific impulse response functions of real stock prices to a one standard deviation increase in oil price uncertainty. The median impulse response is presented in solid lines, while the 16–84% lower and upper bootstrapped error bands are shown in dotted lines. The impact is measured in percentage points (by multiplying the estimates in the figure with 100) over the quarterly horizon.

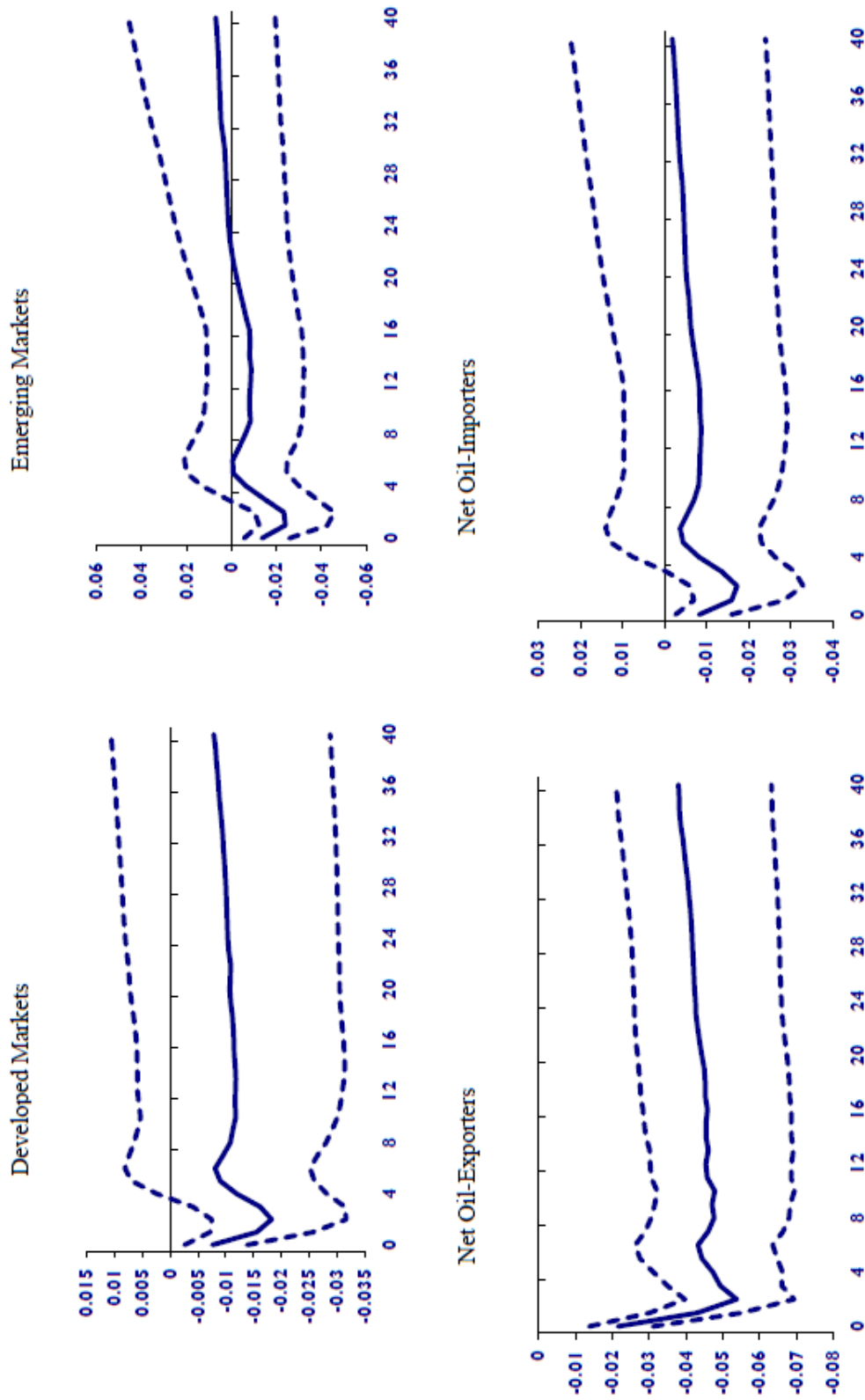


Figure 2. GVAR-based oil price uncertainty shocks and real stock prices (country groups). Note: The figure depicts the impulse response functions of real stock prices of country groups to a one standard deviation increase in oil price uncertainty. The median impulse response is presented in solid lines, while the 16–84% lower and upper bootstrapped error bands are shown in dotted lines. The impact is measured in percentage points (by multiplying the estimates in the figure with 100) over the quarterly horizon.

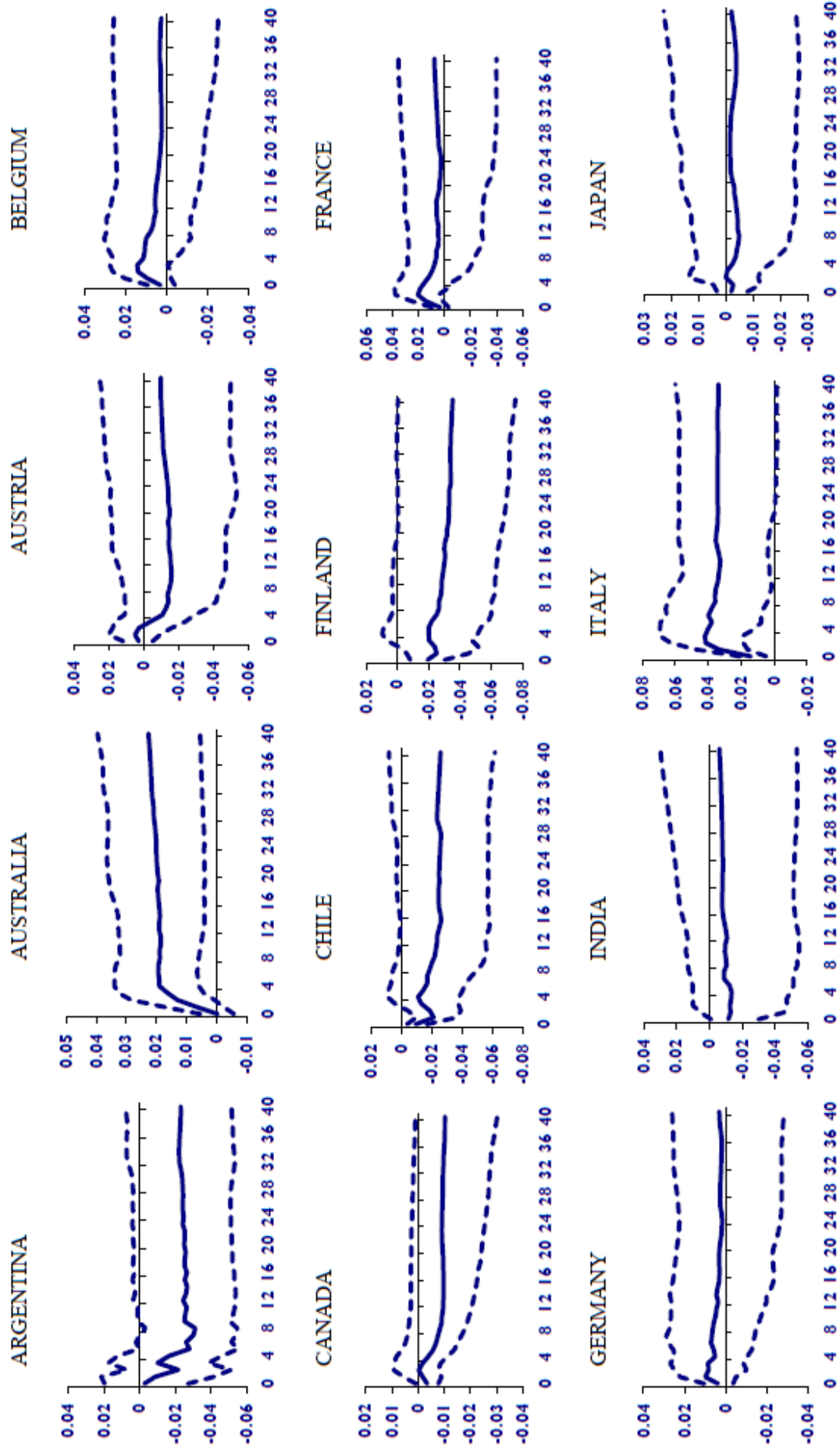


Figure 3. Cont.

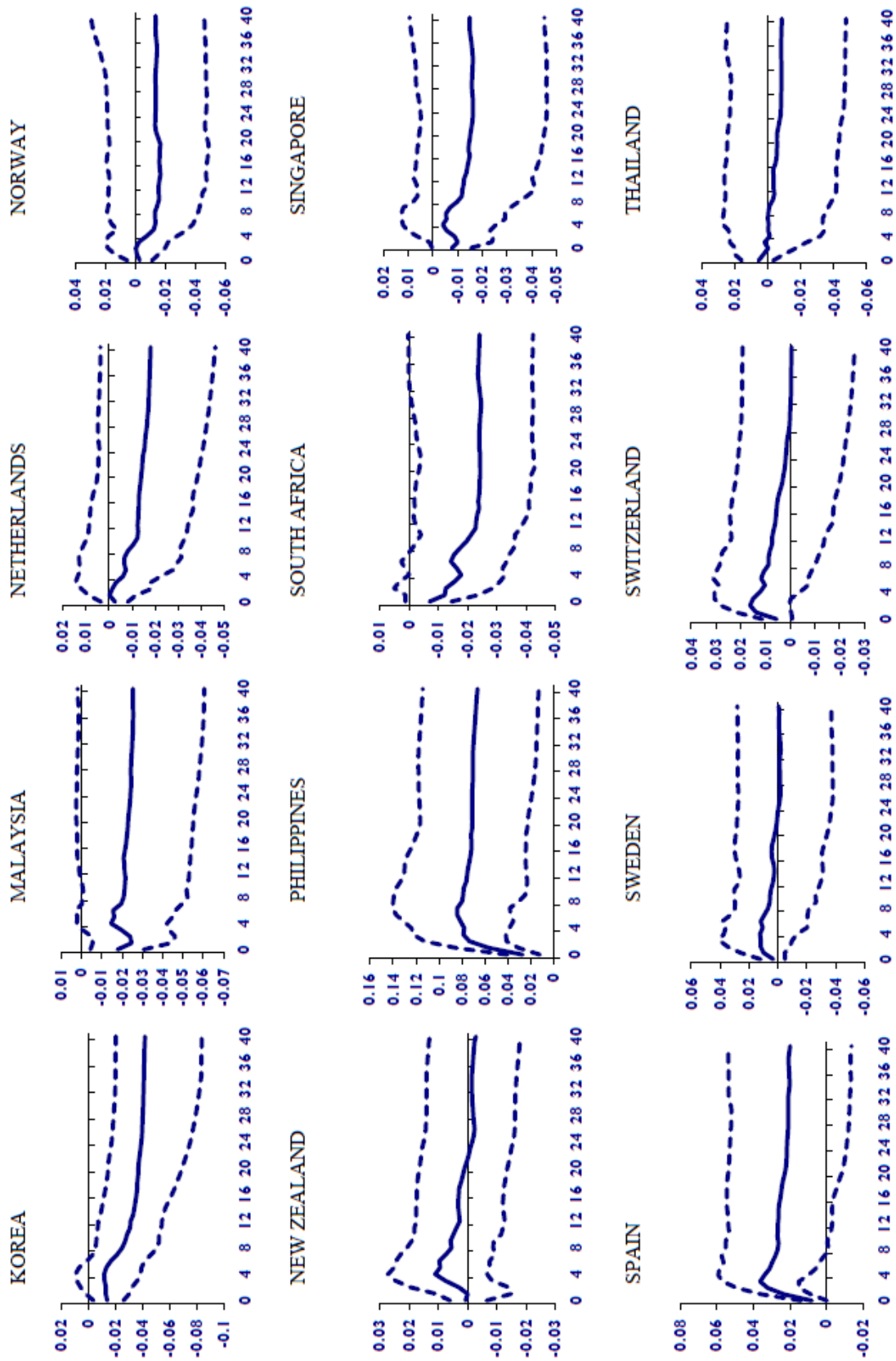


Figure 3. *Cont.*

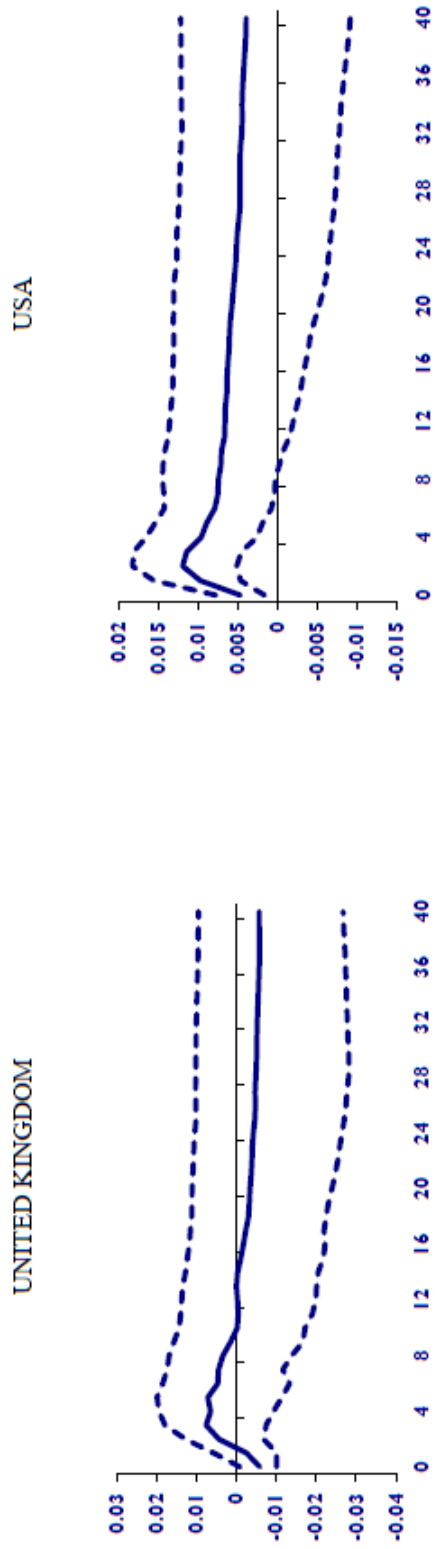


Figure 3. Country-specific oil price uncertainty shocks and real stock prices (*pre-gfc*). Note: The figure depicts the country-specific impulse response functions of real stock prices to a one standard deviation increase in oil price uncertainty. The median impulse response is presented in solid lines, while the 16–84% lower and upper bootstrapped error bands are shown in dotted lines. The impact is measured in percentage points (by multiplying the estimates in the figure with 100) over the quarterly horizon.

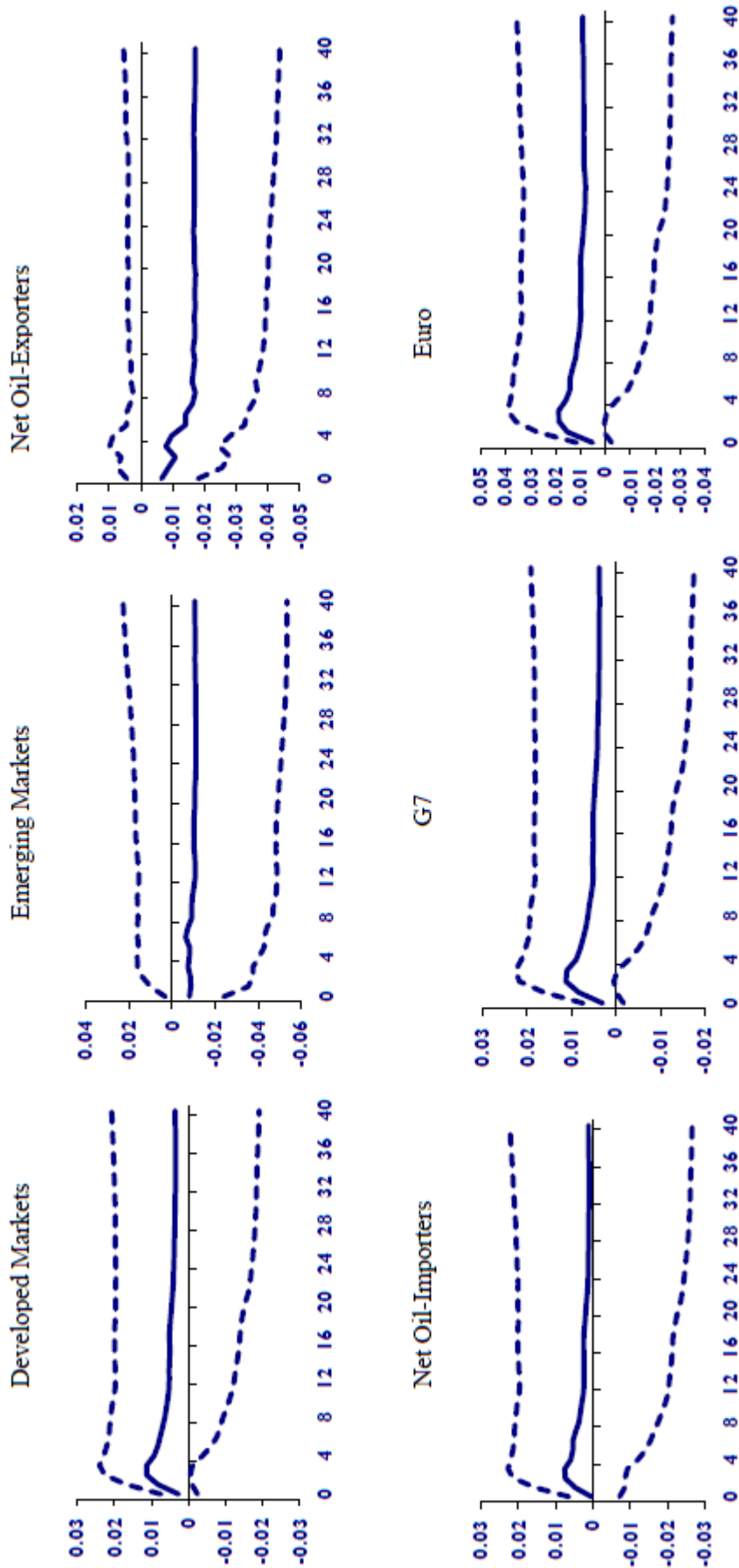


Figure 4. Country group oil price uncertainty shocks and real stock prices (*pre-gfc*). Note: The figure depicts the impulse response functions of real stock prices of country groups to a one standard deviation increase in oil price uncertainty. The median impulse response is presented in solid lines, while the 16–84% lower and upper bootstrapped error bands are shown in dotted lines. The impact is measured in percentage points (by multiplying the estimates in the figure with 100) over the quarterly horizon.

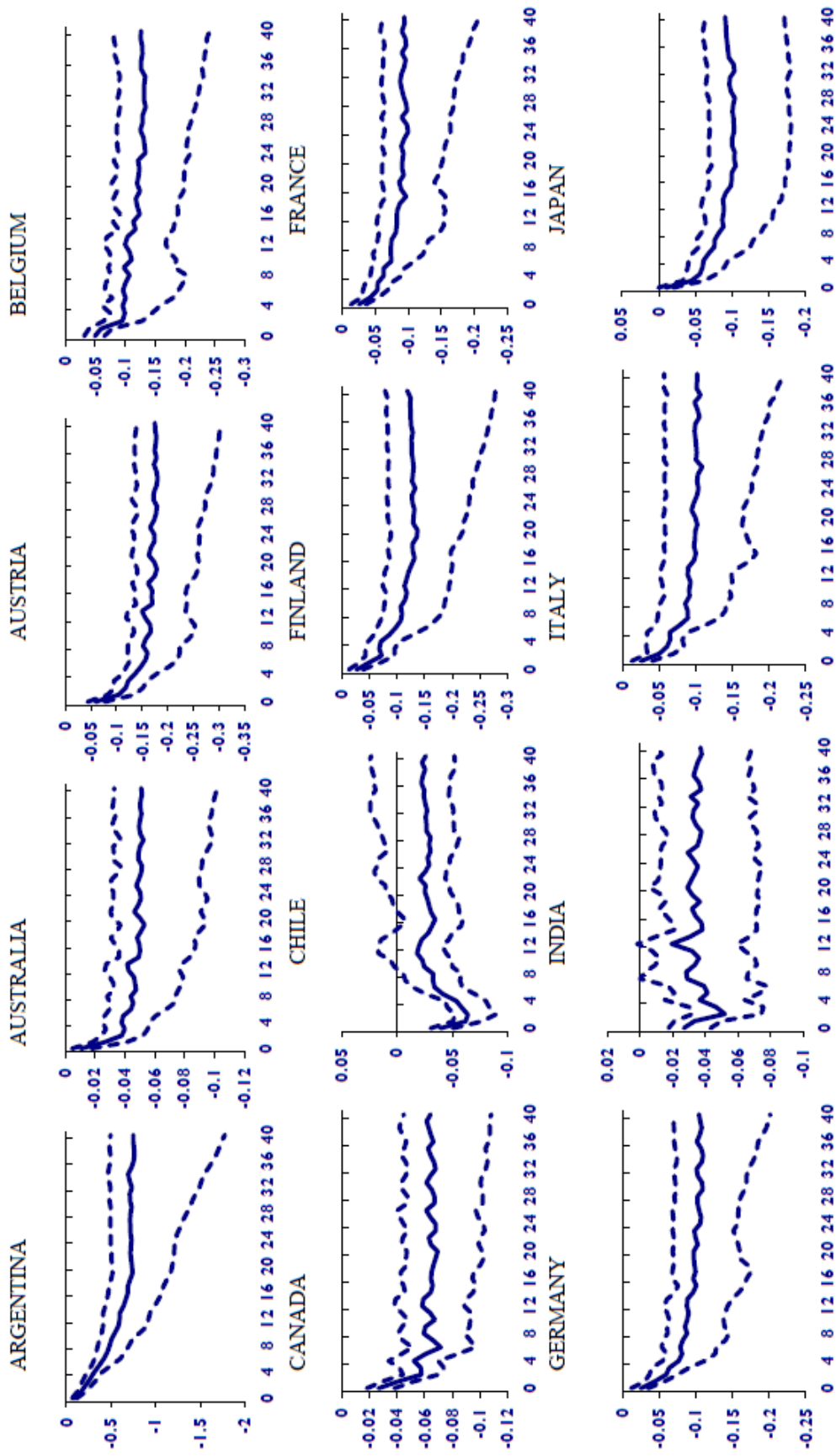


Figure 5. Cont.

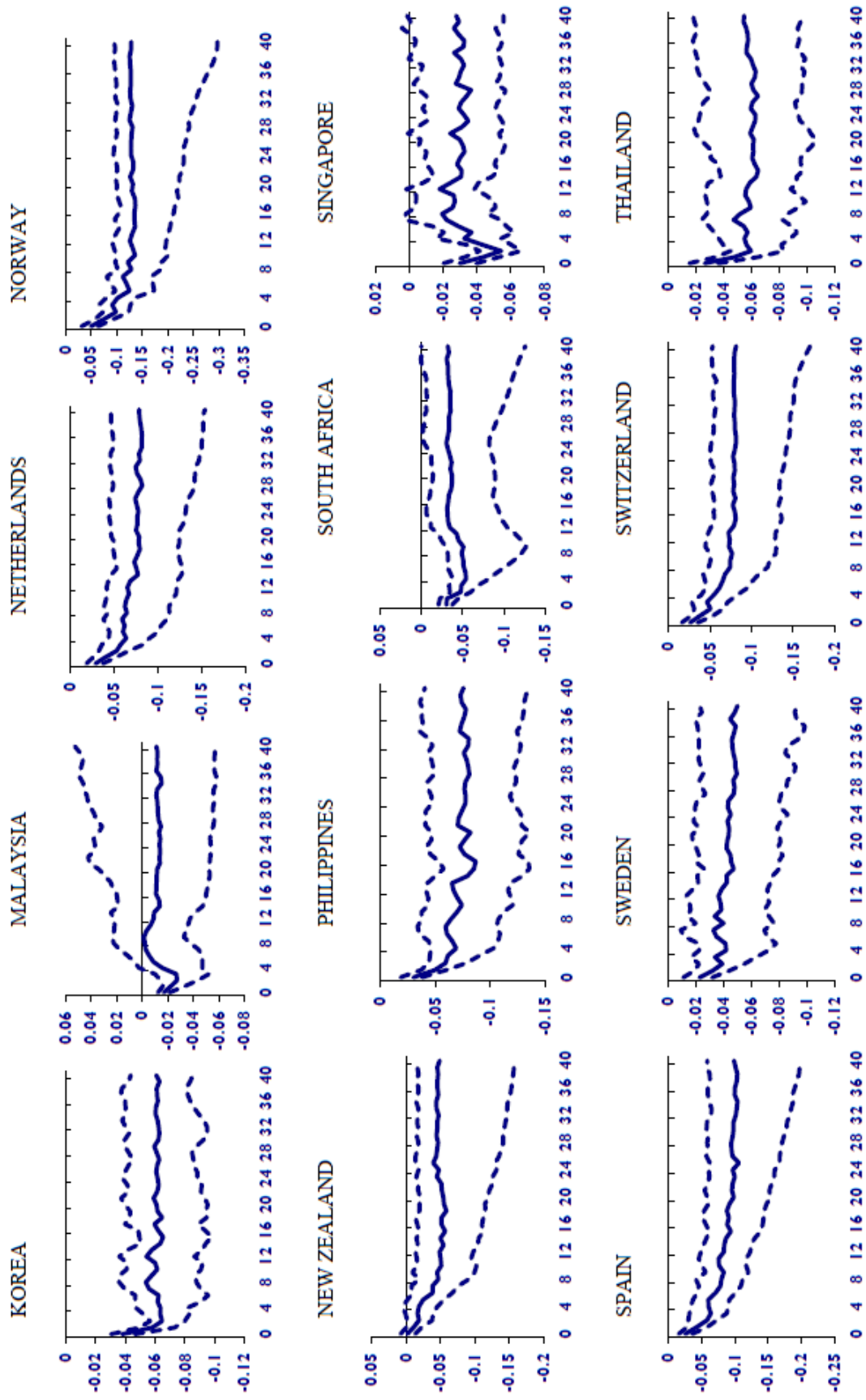


Figure 5. Cont.

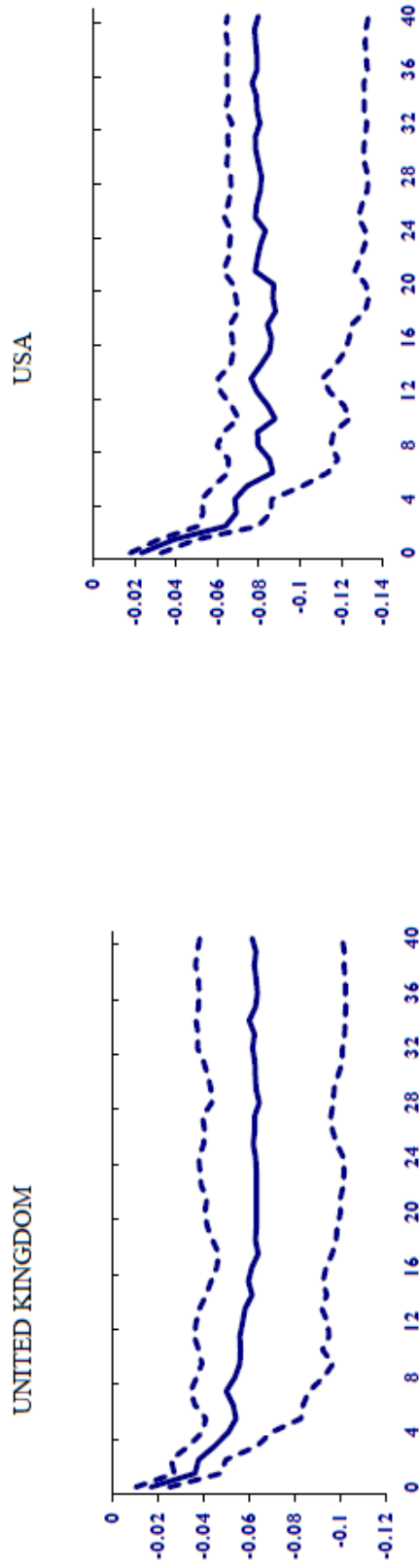


Figure 5. Country-specific oil price uncertainty shocks and real stock prices (*post-gfc*). Note: The figure depicts the country-specific impulse response functions of real stock prices to a one standard deviation increase in oil price uncertainty. The median impulse response is presented in solid lines, while the 16–84% lower and upper bootstrapped error bands are shown in dotted lines. The impact is measured in percentage points (by multiplying the estimates in the figure with 100) over the quarterly horizon.

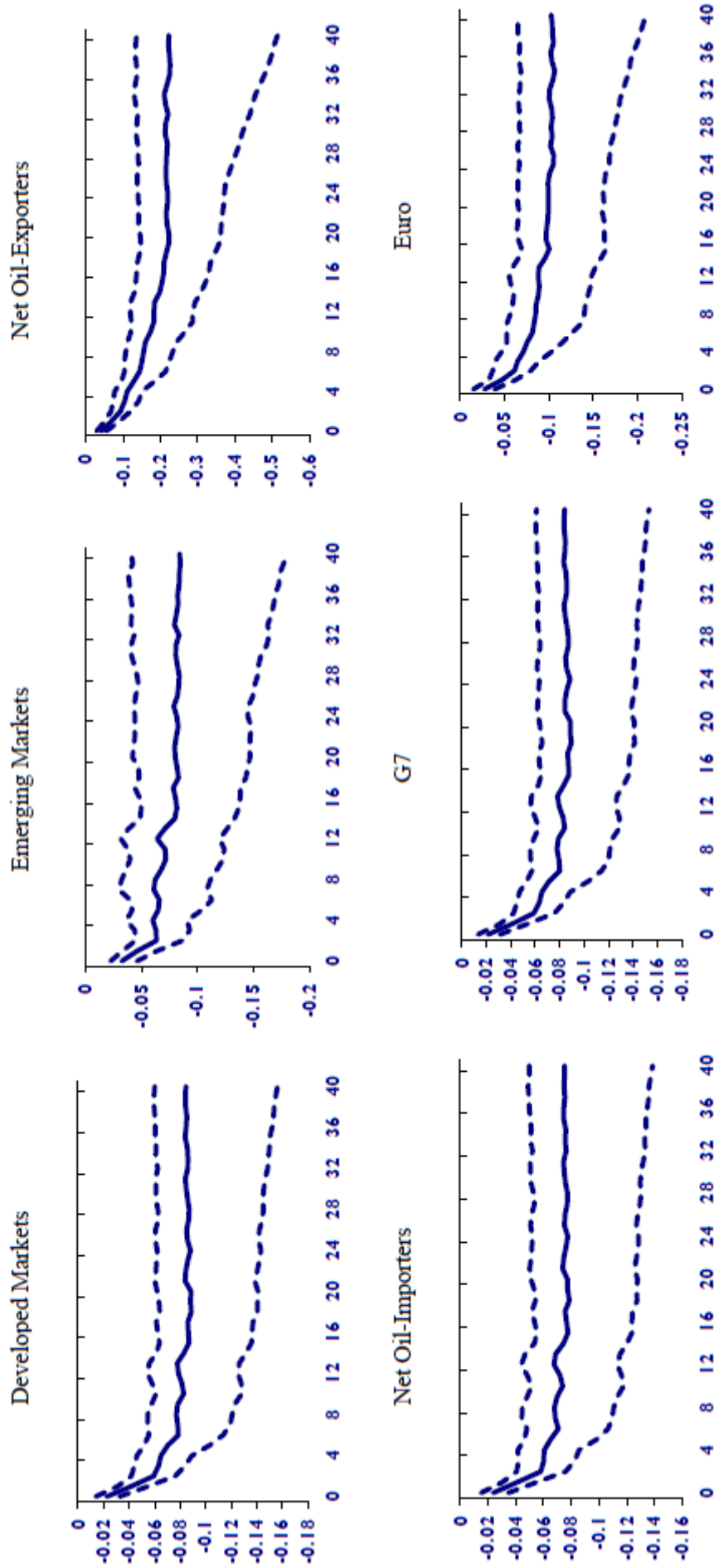


Figure 6. Country group oil price uncertainty shocks and real stock prices (*post-gfc*). Note: The figure depicts the impulse response functions of real stock prices of country groups to a one standard deviation increase in oil price uncertainty. The median impulse response is presented in solid lines, while the 16–84% lower and upper bootstrapped error bands are shown in dotted lines. The impact is measured in percentage points (by multiplying the estimates in the figure with 100) over the quarterly horizon.

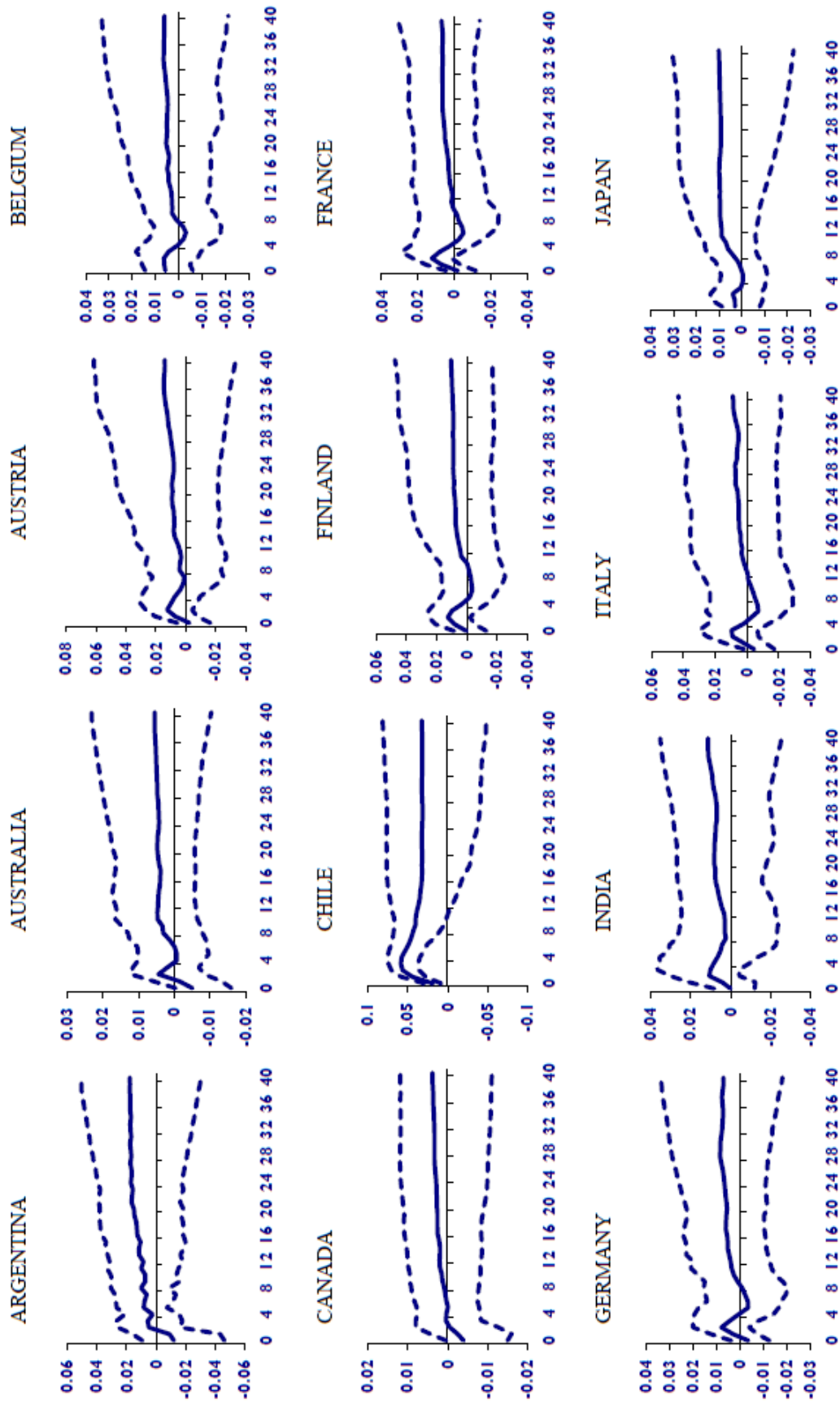


Figure 7. Cont.

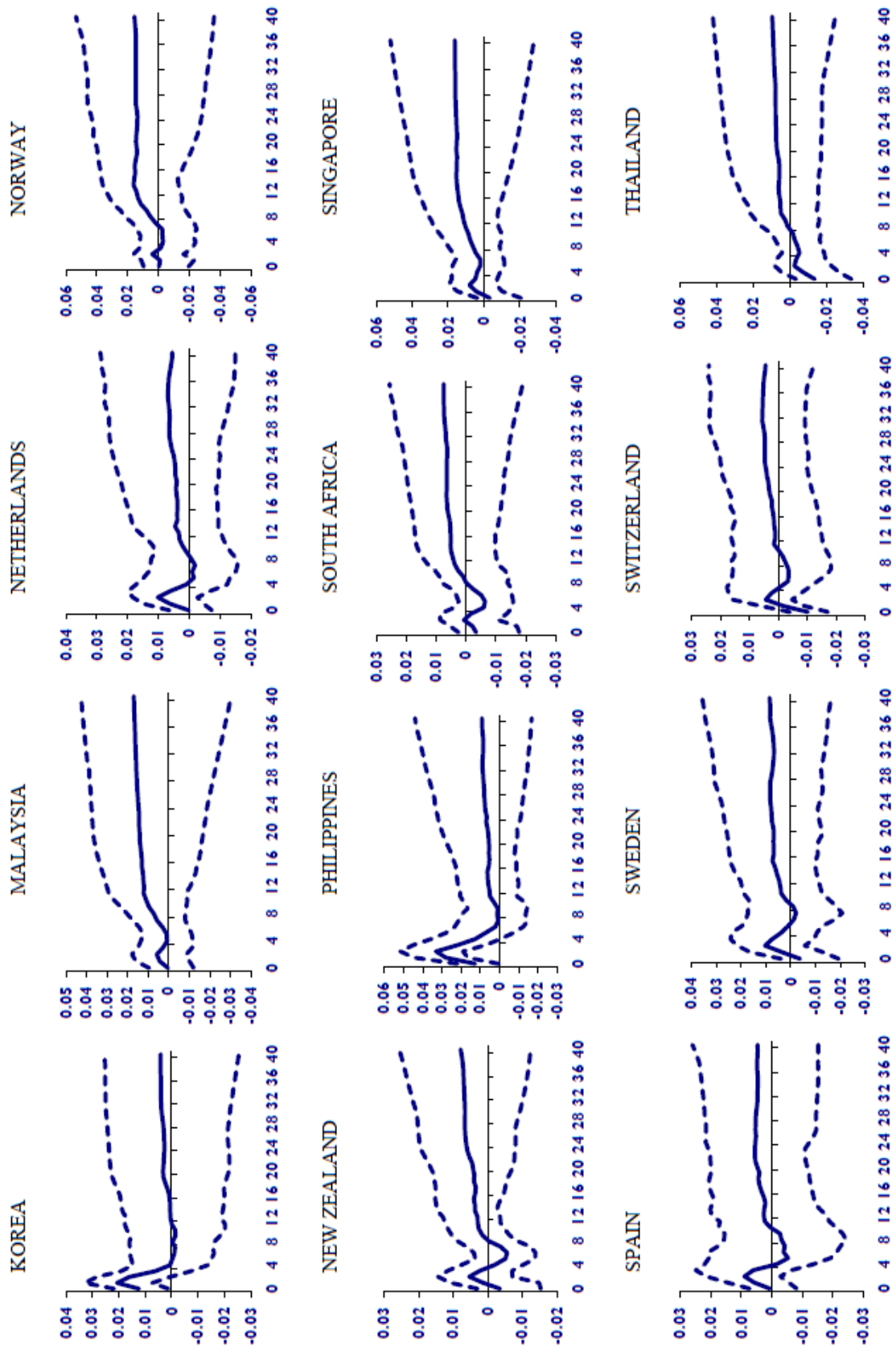


Figure 7. Cont.

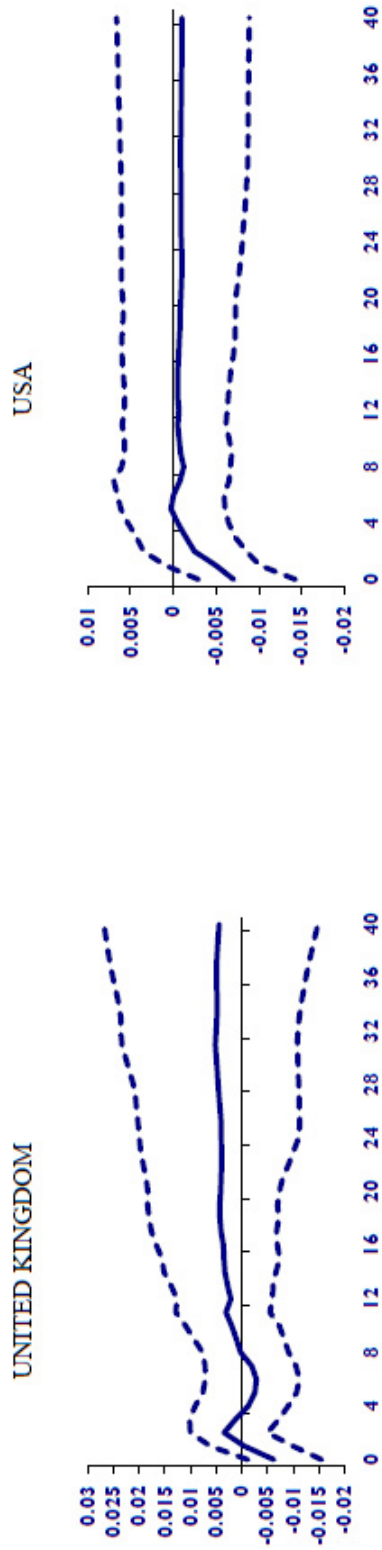


Figure 7. GARCH-based oil price uncertainty shocks and real stock prices (*country-specific*). Note: The figure depicts the country-specific impulse response functions of real stock prices to a one standard deviation increase in oil price uncertainty. The median impulse response is presented in solid lines, while the 16–84% lower and upper bootstrapped error bands are shown in dotted lines. The impact is measured in percentage points (by multiplying the estimates in the figure by 100) over the quarterly horizon.

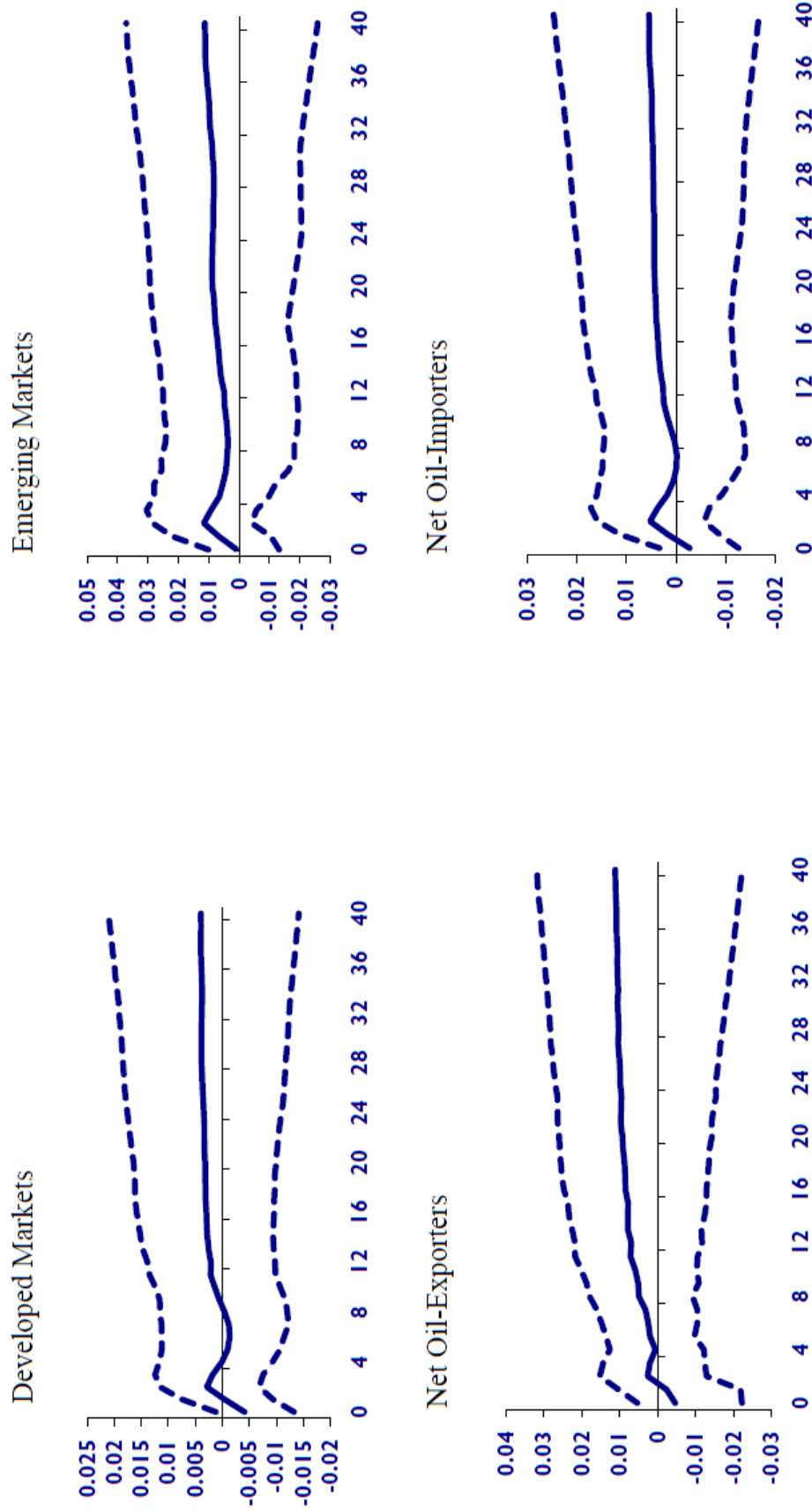


Figure 8. GARCH-based oil price uncertainty shocks and real stock prices (country groups). Note: The figure depicts the impulse response functions of real stock prices of country groups to a one standard deviation increase in oil price uncertainty. The median impulse response is presented in solid lines, while the 16–84% lower and upper bootstrapped error bands are shown in dotted lines. The impact is measured in percentage points (by multiplying the estimates in the figure by 100) over the quarterly horizon.

5. Conclusions

In this study, we examine the nexus between oil and stock markets from a novel perspective by utilising a recently proposed, model-free measure of oil price uncertainty within a Global Vector Autoregressive (GVAR) framework covering 26 global stock markets. We show that oil price uncertainty shocks dampen real stock prices in the large majority of the countries in our sample, with the effect found to be more persistent for those with higher economic complexity and greater reliance on oil in their exports. While the effect of the uncertainty shock is negative for most countries in the short run, we observe a great deal of heterogeneity in the recovery of certain countries/regions from the shock. This is in line with the evidence in the literature that the oil–stock market relationship is not homogeneous across global markets, which offers useful pointers for possible hedging strategies against rising oil market uncertainty. One possible strategy is to devise a conditional investment scheme in which investors hold long positions in emerging market exchange-traded funds (ETFs) funded by short positions in developed stock market ETFs, conditioned on the state of the oil price uncertainty, along the lines of Demirer et al. (2020a).

From a policy perspective, considering that stock prices serve as a leading indicator of macroeconomic variables (Stock and Watson 2003), the impact of oil price uncertainty on equity markets is likely to prolong the direct effects of the same on economic activities (Van Eyden et al. 2019). In other words, high oil price uncertainty depresses economic activities (Salisu et al. 2021), and this could drive firms to delay investment decisions until the uncertainty is lower (Elder and Serletis 2009; Henriques and Sadorsky 2011). Thus, the decision to cut investment usually has an overarching effect on firms' profitability and their stock prices (Bayrakdaroglu et al. 2017). Hence, authorities in oil-rich, as well as emerging economies, should be devising policies to make domestic risk management instruments readily available for firms in the wake of rising oil price uncertainty. This could be in the form of facilitating the development of national derivatives markets in which local firms can utilise futures and options contracts to mitigate their exposure to oil price shocks. For future research, considering the growing evidence that disentangling oil uncertainty based on the nature and origin of the oil price shock in terms of demand or supply can improve the inferences regarding the effect of oil price uncertainty on financial markets (e.g., Demirer et al. 2020b), it would be interesting to implement the disentangled oil uncertainty series within the GVAR framework, which could offer better insights into the heterogeneous responses of stock markets to oil price uncertainty shocks. This is an area we reserve for future research, besides analysing second-moment effects of oil uncertainty on other financial markets as in Liu et al. (2013).

Author Contributions: A.A.S.: Conceptualization, Methodology, Formal Analysis, Writing—Original Draft, Reviewing & Editing. R.G.: Conceptualization, Methodology, Software, Formal Analysis, Writing—Original and Revised Drafts, Reviewing and Editing. R.D.: Conceptualization, Writing—Original and Revised Drafts, Reviewing & Editing, Project Administration. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The GVAR data used for this study can be obtained from <https://sites.google.com/site/gvarmodelling/data>, accessed on 17 July 2022. Data for the oil uncertainty index which is not captured in the GVAR dataset can be obtained from <https://sites.google.com/site/nguyenhoaibao/oil-market-uncertainty?authuser=0>, accessed on 17 July 2022.

Acknowledgments: We would like to thank three anonymous referees for many helpful comments. However, any remaining errors are solely ours.

Conflicts of Interest: The authors declare no conflict of interest.

Notes

- ¹ Basher and Sadorsky (2006) employ a multi-factor model to show that oil price risk tends to strongly affect a large set of emerging stock market returns.
- ² The developed markets include Australia, Austria, Belgium, Canada, Finland, France, Germany, Italy, Japan, The Netherlands, Norway, New Zealand, Singapore, Spain, Sweden, Switzerland, United Kingdom and the USA; while the emerging markets are Argentina, Chile, India, Malaysia, Philippines, South Africa, South Korea and Thailand. Our choice of countries is guided by data availability, as we consider a more representative classification for all the countries following the Morgan Stanley Classification Index (MSCI). In other words, countries with fewer number of years as being developed but more years as emerging are classified as emerging, and same holds for our consideration for the developed countries.
- ³ <https://www.econ.cam.ac.uk/people-files/emeritus/mhp1/GVAR/GVAR.html>, accessed on 17 July 2022.
- ⁴ Data is available at <https://www.econ.cam.ac.uk/people-files/emeritus/mhp1/GVAR/GVAR.html>, accessed on 17 July 2022. However, data for the oil uncertainty index is not captured in the GVAR dataset and can be obtained from <https://sites.google.com/site/nguyenhoaibao/oil-market-uncertainty?authuser=0>, accessed on 17 July 2022. Note that, the monthly values of the oil price uncertainty index is converted to quarterly values by taking three-month averages over a quarter to match the quarterly frequency of the GVAR dataset.
- ⁵ The oil price uncertainty index is based on the conditional volatility of the unpredictable component of the real price of oil as measured by the Consumer Price Index (CPI)-deflated nominal values of the conventional US refiners' acquisition cost for imported crude oil. The reader is referred to Nguyen et al. (2021) for further technical details.
- ⁶ See the link for the ranking and discussion of economic complexity <https://www.visualcapitalist.com/countries-ranked-by-their-economic-complexity/>, accessed on 17 July 2022. Highly ranked countries in terms of economic complexity imply a high diversity of exported products and sophisticated and unique exported products (i.e., few other countries produce similar products).
- ⁷ See the link for the ranking of countries by oil consumption per capita <https://www.eia.gov/tools/faqs/faq.php?id=709&t=6>, accessed on 17 July 2022.
- ⁸ The classification of countries into emerging and developed economies is drawn from the market classification by the Morgan Stanley Capital International (MSCI) (see <https://www.msci.com/market-classification>, accessed on 17 July 2022), while the World Fact Book of the Central Intelligence Agency (CIA) is used to group countries into net oil exporting and net oil importing countries.

References

- Alsalmán, Zeina. 2016. Oil price uncertainty and the US stock market analysis based on a GARCH-in-mean VAR model. *Energy Economics* 59: 251–60. [CrossRef]
- Aye, Goodness. 2015. Does oil price uncertainty matter for stock returns in South Africa? *Investment Management and Financial Innovations* 12: 179–88.
- Aye, Goodness, Mehmet Balçilar, and Rangan Gupta. 2017. International stock return predictability: Is the role of US time-varying? *Empirica* 44: 121–46. [CrossRef]
- Basher, Syed Abul, Alfred Haug, and Perry Sadorsky. 2012. Oil prices, exchange rates and emerging stock markets. *Energy Economics* 34: 227–40. [CrossRef]
- Basher, Syed Abul, and Perry Sadorsky. 2006. Oil price risk and emerging stock markets. *Global Finance Journal* 17: 224–51. [CrossRef]
- Bass, Alexander. 2017. Does oil prices uncertainty affect stock returns in Russia: A bivariate GARCH-in-mean approach. *International Journal of Energy Economics and Policy* 7: 224–30.
- Bayrakdaroglu, Ali, Çağatay Mirgen, and Ezgi Kuyu. 2017. Relationship between profitability ratios and stock prices: An empirical analysis on BIST-100. *Press Academia Procedia* 6: 1–10. [CrossRef]
- Benavides, Domingo Rodríguez, Miguel Ángel Martínez García, and Luis Fernando Hoyos Reyes. 2019. Uncertainty of the international oil price and stock returns in Mexico through an SVAR-MGARCH. *Contaduría y administración* 64: 1–16.
- Bernanke, Ben Shalom. 1983. Irreversibility, uncertainty, and cyclical investment. *Quarterly Journal of Economics* 98: 85–106. [CrossRef]
- Chen, Chun-Da, and Riza Demirel. 2022. Oil beta uncertainty and global stock returns. *Energy Economics* 112: 106150. [CrossRef]
- Chen, Shiu-Sheng. 2009. Do higher oil prices push the stock market into bear territory? *Energy Economics* 32: 490–95. [CrossRef]
- Chudik, Alexander, and Mohammad Hashem Pesaran. 2016. Theory and practice of GVAR modelling. *Journal of Economic Surveys* 30: 165–97. [CrossRef]
- Demirel, Riza, Aydin Yuksel, and Asli Yuksel. 2020a. Oil price uncertainty, global industry returns and active investment strategies. *Journal of Economic Asymmetries* 22: e00177. [CrossRef]
- Demirel, Riza, Rangan Gupta, Christian Pierdzioch, and Syed Jawad Hussain Shahzad. 2020b. The predictive power of oil price shocks on realized volatility of oil: A note. *Resources Policy* 69: 101856. [CrossRef] [PubMed]
- Demirel, Riza, Shrikant Jategaonkar, and Ahmed Khalifa. 2015. Oil price risk exposure and the cross-section of stock returns: The case of net exporting countries. *Energy Economics* 49: 142–40. [CrossRef]
- Diaz, Elena Maria, Juan Carlos Molero, and Fernando Perez de Gracia. 2016. Oil price volatility and stock returns in the G7 economies. *Energy Economics* 54: 417–30. [CrossRef]

- Elder, John, and Apostolos Serletis. 2009. Oil price uncertainty in Canada. *Energy Economics* 31: 852–56. [CrossRef]
- Hatemi, J. Abdalnasser, Abdulrahman Al Shayeb, and Eduardo Roca. 2017. The effect of oil prices on stock prices: Fresh evidence from asymmetric causality tests. *Applied Economics* 49: 1584–92. [CrossRef]
- Henriques, Irene, and Perry Sadorsky. 2011. The effect of oil price volatility on strategic investment. *Energy Economics* 33: 79–87. [CrossRef]
- Jiranyakul, Komain. 2014. Does oil price uncertainty transmit to the Thai stock market? *Journal of Economic & Financial Studies* 2: 16–25.
- Jo, Soojin. 2014. The effects of oil price uncertainty on global real economic activity. *Journal of Money, Credit and Banking* 46: 1113–35. [CrossRef]
- Jordan, Steven, Andrew Vivian, and Mark Wohar. 2016. Forecasting market returns: Bagging or combining? *International Journal of Forecasting* 33: 102–20. [CrossRef]
- Jordan, Steven, Andrew Vivian, and Mark Wohar. 2017. Stock returns forecasting with metals: Sentiment vs. fundamentals. *The European Journal of Finance* 24: 458–77. [CrossRef]
- Liu, Ming-Lei, Qiang Ji, and Ying Fan. 2013. How does oil market uncertainty interact with other markets? An empirical analysis of implied volatility index. *Energy* 55: 860–68. [CrossRef]
- Maghyereh, Aktham, and Hussein Abdoh. 2020. Asymmetric effects of oil price uncertainty on corporate investment. *Energy Economics* 86: 1–19. [CrossRef]
- Masih, Rumi, Sanjay Peters, and Lurion De Mello. 2011. Oil price volatility and stock price fluctuations in an emerging market: Evidence from South Korea. *Energy Economics* 33: 975–86. [CrossRef]
- Mohaddes, Kamiar, and Mehdi Raissi. 2020. *Compilation, Revision and Updating of the Global VAR (GVAR) Database, 1979Q2-2019Q4*. Mimeo: University of Cambridge, Judge Business School. [CrossRef]
- Nguyen, Bao, Tatsuyoshi Okimoto, and Trung Duc Tran. 2021. Uncertainty-dependent and sign-dependent effects of oil market shocks. *Journal of Commodity Markets* 26: 100207. [CrossRef]
- Pesaran, Mohammad. Hashem, Til Schuermann, and Scott Weiner. 2004. Modeling Regional Interdependencies Using a Global Error-Correcting Macroeconometric Model. *Journal of Business and Economic Statistics* 22: 129–62. [CrossRef]
- Pindyck, Robert. 1991. Irreversibility, uncertainty, and investment. *Journal of Economic Literature* 29: 1110–48.
- Prodromou, Tina, and Riza Demirer. 2022. Oil Price Shocks and Cost of Capital: Does Market Liquidity Play a Role? Available online: <https://ssrn.com/abstract=4144883> (accessed on 17 July 2022).
- Rahman, Sajjadur. 2021. Oil price volatility and the US stock market. *Empirical Economics* 61: 1461–89. [CrossRef]
- Sadorsky, Perry. 1999. Oil price shocks and stock market activity. *Energy Economics* 21: 449–69. [CrossRef]
- Salisu, Afees Adebare, Rangan Gupta, and Abee Olaniyan. 2021. The effect of oil uncertainty shock on real GDP of 33 countries: A global VAR approach. *Applied Economics Letters* 1–6. [CrossRef]
- Salisu, Afees Adebare, and Kazeem Isah. 2017. Revisiting the oil price and stock market nexus: A nonlinear Panel ARDL approach. *Economic Modelling* 66: 258–71. [CrossRef]
- Silvapulle, Param, Russell Smyth, Xibin Zhang, and Jean-Pierre Fenech. 2017. Nonparametric panel data model for crude oil and stock prices in net oil importing countries. *Energy Economics* 67: 255–67. [CrossRef]
- Sousa, Ricardo, Andrew Vivian, and Mark Wohar. 2016. Predicting asset returns in the BRICs: The role of macroeconomic and fundamental predictors. *International Review of Economics and Finance* 41: 122–43. [CrossRef]
- Stock, James, and Mark Watson. 2003. Forecasting Output and Inflation: The Role of Asset Prices. *Journal of Economic Literature* XLI: 788–829.
- Swaray, Raymond, and Afees Adebare Salisu. 2018. A firm-level analysis of the upstream-downstream dichotomy in the oil-stock nexus. *Global Finance Journal* 37: 199–218. [CrossRef]
- Van Eyden, Renee, Mamothoana Difeto, Rangan Gupta, and Mark Wohar. 2019. Oil price volatility and economic growth: Evidence from advanced economies using more than a century's data. *Applied Energy* 233: 612–21. [CrossRef]
- Wang, Yudong, Chongfeng Wu, and Li Yang. 2013. Oil price shocks and stock market activities: Evidence from oil-importing and oil-exporting countries. *Journal of Comparative Economics* 41: 1220–39. [CrossRef]
- Yin, Libo, and Man Lu. 2022. Oil uncertainty and firms' risk-taking. *Energy Economics* 108: 105922. [CrossRef]
- Zhu, Hui-Ming, Su-Fang Li, and Keming Yu. 2011. Crude oil shocks and stock markets: A panel threshold cointegration approach. *Energy Economics* 33: 987–94. [CrossRef]

Article

Dynamic Relationship between Volatility Risk Premia of Stock and Oil Returns

Nobuhiro Nakamura , Kazuhiko Ohashi * and Daisuke Yokouchi

Graduate School of Business Administration, Hitotsubashi University, 2-1-2 Hitotsubashi, Chiyoda-ku, Tokyo 101-8439, Japan

* Correspondence: kohashi@hub.hit-u.ac.jp

Abstract: This study investigates the relationship between the volatility risk premia (VRP) of stock and oil returns. Using daily data on VRP from 10 May 2007 to 16 May 2017, VAR analyses on the stock and oil VRP are conducted, and it is found that the effects of the stock VRP on the oil VRP are limited and, if any, short-lived. In contrast, the VRP of oil has significantly positive and long-lasting effects on the stock VRP after the financial crisis. These results suggest that investors' sentiments (measured by VRP) are transmitted from the oil to the stock market over time, but not vice versa. This is unexpected because the financialization of commodities means a massive increase in investment in commodities by investors in the traditional stock and bond markets; hence, the direction of effects is thought to be from the stock to the commodity market.

Keywords: volatility risk premium (VRP); implied and realized volatility; oil and stock returns; financialization

JEL Classification: G11; G12; G13



Citation: Nakamura, Nobuhiro, Kazuhiko Ohashi, and Daisuke Yokouchi. 2023. Dynamic Relationship between Volatility Risk Premia of Stock and Oil Returns. *Journal of Risk and Financial Management* 16: 173. <https://doi.org/10.3390/jrfm16030173>

Academic Editor: Kentaro Iwatsubo

Received: 27 December 2022

Revised: 17 February 2023

Accepted: 27 February 2023

Published: 5 March 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The volatility risk premium (VRP), defined as the difference between implied and realized volatilities, has been found to have predictive power for returns in many different assets. For example, as pioneering research on this topic, Bollerslev et al. (2009a) revealed that VRP has predictive power for U.S. monthly aggregated stock returns, and Bollerslev et al. (2009b) also found the predictive power of the VRP in the monthly stock index returns of many other developed countries.

The VRP represents the risk premium for future volatility variations. Thus, it may be regarded as investor sentiment (i.e., aversion to future uncertainty), and the predictability of VRP is thought to be due to investor sentiment: when investor sentiment worsens (resp. improves), stock prices are discounted by a higher (resp. lower) premium, resulting in higher (resp. lower) future returns.

Following this intuition, the scope of the analysis is extended to assets other than stocks. Indeed, Della Corte et al. (2016) and Londono and Zhou (2016) confirmed the predictive power of VRP in monthly exchange rates. Furthermore, Ornelas and Mauad (2019) investigated the predictive power of different assets' VRP such as commodities currencies, stocks, bonds, gold, and oil, on the monthly returns.

Given the extant research on VRP's return predictability of different assets, one simple but unexplored question is how the VRP of different assets are correlated. This question is meaningful because the dynamic relation of the VRP between different assets is interpreted to show how investors' sentiments on different assets transmit to each other over time. Moreover, it is especially important between the VRP of stocks and commodities because the recent financialization of commodities, that is the massive increase in investment in commodities by investors in the traditional stock and bond markets, is thought to increase

the influence of the stock market on commodity markets. Thus, this study investigates the dynamic relationship between the VRP of stocks and oil using their daily returns.

Note that this focus is unique compared to previous studies. Indeed, many extant studies have investigated the relationship among implied volatilities of stock, oil, gold, and exchange rate. However, they are not related to VRP. For example, Robe and Wallen (2016) analyzed the determinants of oil implied volatility using weekly data and investigated the relationship between oil implied volatility and stock implied volatility. Christoffersen and Pan (2018) investigated the effect of oil implied volatility on stock returns and analyzed the relationship between oil implied volatility and stock implied volatility. Liu et al. (2013) conducted a VECM analysis on the relation among daily implied volatilities of stock, oil, gold, and exchange rates. Dutta et al. (2019) analyzed co-integration and nonlinear causality among the implied volatilities of crude oil, gold, silver and goldminers by a non-linear ARDL model. Bouri et al. (2020) studied the dynamic spillovers among the implied volatilities of the S&P 500 and five large US stocks based on Diebold and Yilmaz's (2014) connectedness model. Iqbal et al. (2022) analyzed spillover among implied volatilities of international stock and commodity indices by a quantile VAR model. Moreover, Gagnon et al. (2015), Zhang et al. (2022), and Bouri et al. (2023) analyzed the relationship among implied higher order moments of stock indices and commodities.

Given those previous studies, the main goal of this study is to investigate the dynamic relation of daily VRP, not implied volatilities, between stocks and oil, and to show how investors' sentiments, represented by VRP, on different assets transmit each other over time. A paper closely related to this is Hattori et al. (2021), who conducted a VAR analysis on the relationship among daily stock VRP of advanced and emerging market economies. Our study differs in that it investigates the dynamic relationship of VRP between stock and oil and analyzes the spillover of investors' sentiments, not within stock markets, but between stock and commodity markets. To the best of our knowledge, this is the first study to address this issue.

Following the method of Bollerslev et al. (2009a), we calculate the daily VRP of stock as the difference between the VIX published by the Chicago Board of Trade (CBOE), which measures the 30-day implied volatility of S&P 500 stock index options, and the daily realized volatility of the S&P 500 stock index provided by the Oxford-Man Institute of Quantitative Finance. To obtain the daily VRP of oil, we use the OVX published by the CBOE, which measures the 30-day implied volatility of crude oil prices by applying the CBOE Volatility Index methodology to options on the United States Oil Fund (USO). Because we do not have high-frequency data of the USO prices and hence cannot directly calculate its daily realized volatility, we estimate the daily realized volatility of oil by applying a stochastic volatility model to its returns (see Appendix A).

Using the daily VRP of stock and oil returns obtained between 10 May 2007 and 16 May 2017, we conduct a VAR analysis of the VRP and obtain the following results: During the whole period and all sub-periods, both VRP are stationary and their correlations are approximately 0.2 to 0.3, except in the pre-crisis period (between 10 May 2007 and 30 May 2008), where the correlation is less than 0.1.

For the whole period, most of the variations in the VRP are explained by their own shocks, which may seem against what we expect from the financialization of commodities because financialization is regarded as strengthening the relationship between stock and oil. Meanwhile, the shocks in both the VRP of stock and oil have small but significant positive effects on each other for most of the following 20 trading days after the shock. This is in contrast with the results shown by Liu et al. (2013) on the relationships among the implied volatilities of stock, oil, gold, and euro/dollar exchange rates, in which all implied volatilities have significant, but only temporary (i.e., just on the 1st trading day after the shock) effects on each other.

However, such relationships depend on the economic situation, and the economic situation surrounding stock and oil has been clearly changing. Thus, we conduct a VAR analysis on the following sub-periods: Period 1 from 10 May 2007 to 31 May 2008 (pre-crisis

period), Period 2 from 1 June 2008 to 30 June 2009 (crisis outbreak period), Period 3 from 1 July 2009 to 31 July 2012 (post-crisis recovery period I), Period 4 from 1 August 2012 to 30 September 2014 (post-crisis recovery period II), and Period 5 from 1 October 2014 to 16 May 2017 (plunging oil price period). Interestingly, the analysis of these sub-periods reveals a different picture of the dynamic relationship between stock and oil VRP from that of the entire period.

In the pre-crisis period (Period 1), we find that there is little or no relation between the VRP of stock and oil: a small correlation of less than 0.1, no Granger causality between the stock and oil VRP, or no significant effects on each other in impulse response functions and little effect on variance decomposition. Again, this may seem somewhat against the view of the financialization of commodities because the financialization effect, that is the rise of correlations among the returns of stock and commodities, emerged after the 2000s (Tang and Xiong 2012; Silvennoinen and Thorp 2013; Ohashi and Okimoto 2016).

In the crisis outbreak period (Period 2), the correlation between the VRP of stock and oil is 0.27. The stock VRP does not Granger cause an oil VRP. There are no significant effects of the stock VRP on the oil VRP in either the impulse response functions or variance decomposition. In contrast, the oil VRP Granger causes the stock VRP, has significantly positive, though small, effects on the stock VRP in impulse response and explains 8% of the variation in the stock VRP.

In post-crisis recovery period I (Period 3), the correlation increases to 0.34. Both the stock and oil VRP Granger cause each other. Both have small but significantly positive effects on each other in the impulse response functions and variance decomposition. However, their effects have quite different patterns: The VRP of oil has significantly positive and long-lasting effects (after the 2nd trading day of the shock), whereas the VRP of stock has significantly positive but only temporary effects (just up to the 2nd trading day) on that of oil.

In post-crisis recovery period II (Period 4), the correlation decreases to 0.21. The Granger causality from stock to oil disappears, while the VRP of oil Granger causes that of the stock. The effects of the oil VRP on the stock VRP remain significant and long-lasting, similar to those in Period 3, but the effects of the stock VRP on the oil VRP disappear.

Finally, in the plunging oil price period (Period 5), the correlation is 0.22. Both Granger cause one another. The effects of the VRP of stock on that of oil are back to significantly positive, but only temporarily on the 1st trading day after the shock, while the effects of the oil VRP on the stock VRP remain significantly positive up to the 8th trading day after the shock.

In summary, the dynamic relationship between the VRP of stock and oil depends on the economic situation, and contrary to the results for the whole period, it is revealed that the VRP of oil has significantly positive and long-lasting effects on that of stock in all sub-periods after the outbreak of the financial crisis, while the effects of the stock VRP on the oil VRP are limited and, if any, much more short-lived. That is, although small, investors' sentiments are transmitted from the oil market to the stock market over time, but not vice versa. This relationship between oil and stock VRP is an unexplored point in the extant literature and is rather unexpected because the financialization of commodities, that is the massive increase in investment in commodities by investors in the traditional stock and bond markets, is thought to have effects from the stock to the commodity market.¹

The remainder of this paper is organized as follows. Section 2 explains VRP. Section 3 discusses the construction and properties of the data used in the study. Section 4 describes the model selection. Section 5 presents our main empirical results. Section 6 discusses the robustness of the analysis. Section 7 provides the conclusion.

2. Volatility Risk Premium (VRP)

Let t denote the current date. Denote by σ_{t+T} the volatility of an asset return at date $t + T$. A volatility swap that exchanges on date $t + T$ the payoff σ_{t+T} and the payment x_t , which is contracted at date t , enables its holder/investor to hedge on date t the risk of

volatility variation in the future date $t + T$.² A simple no-arbitrage argument shows that $x_t = E_t^Q[\sigma_{t+T}]$, where Q is the risk-neutral probability. Hence, the amount that the swap investor receives on date $t + T$ is equal to $\sigma_{t+T} - E_t^Q[\sigma_{t+T}]$.

If σ_{t+T} is on average less than $E_t^Q[\sigma_{t+T}]$, that is, $E_t^P[\sigma_{t+T}] - E_t^Q[\sigma_{t+T}] < 0$ where P is the original probability, it means that the swap holder is willing to pay $E_t^Q[\sigma_{t+T}]$, which is more than the expected payoff $E_t^P[\sigma_{t+T}]$, to hedge the volatility risk in the future. In this sense, $E_t^Q[\sigma_{t+T}] - E_t^P[\sigma_{t+T}]$ represents the premium the swap investor is willing to pay to hedge the variation risk of future volatility. Thus, $E_t^Q[\sigma_{t+T}] - E_t^P[\sigma_{t+T}]$ is the volatility risk premium (VRP).

As visible, the larger the VRP is, the more averse the investor is about the variation in future volatility. In this sense, the VRP is sometimes interpreted as indicating investor sentiment on future asset returns.

3. Data

In the empirical analysis, we estimate the risk-neutral expected future volatility $E_t^Q[\sigma_{t+T}]$ and the expected future volatility $E_t^P[\sigma_{t+T}]$ to calculate the VRP. The former can be estimated from option prices, and hence is called the (option) implied volatility (IV). However, the latter estimation is not immediate. Hence, following the method of Bollerslev et al. (2009a), we approximate the expected future volatility by the realized volatility (RV) and obtain VRP as the difference between IV and RV, that is, $VRP \equiv IV - RV$.³

More precisely, we calculate the daily VRP of stock (VRP_{sp}) as the difference between the VIX published by the Chicago Board of Trade (CBOE),⁴ which measures the 30-day implied volatility of S&P 500 stock index options, and the daily realized volatility of the S&P 500 stock index provided by the Oxford-Man Institute of Quantitative Finance, which is calculated from 5 min returns of the index.⁵ As noted above, while the former is risk-neutral expected future volatility, the latter is not expected future volatility, but daily realized volatility. Thus, by this choice of variable, we assume that the daily realized volatility of the Oxford-Man Institute approximates the expected future volatility of the stock well.

To obtain the daily VRP of oil (VRP_{oil}), we use the OVX published by the CBOE,⁶ which measures the 30-day implied volatility (i.e., the risk-neutral expected future volatility) of crude oil prices by applying the CBOE Volatility Index methodology to options on the United States Oil Fund (USO). Because we do not have high-frequency data of the USO returns to calculate its daily realized volatility, we estimate the daily realized volatility of oil by applying a stochastic volatility model to its returns.⁷ Then, we obtain the daily VRP of oil (VRP_{oil}) as the difference between the OVX and the daily realized volatility of oil. Again, by doing so, we assume that the daily realized volatility of oil approximates the expected future volatility of oil.

As the CBOE publishes OVX data after the middle of 2007, we use the daily VRP of stock and oil returns from 10 May 2007 to 16 May 2017. In this period, however, global financial markets and the world economy went through several different phases such as the global financial crisis around the collapse of Lehman Brothers, the recovery from the financial crisis, and the plunge of oil prices, all of which may affect the relationship between stock and oil prices.

For example, this is visible from graphs of the daily indices of stock (S&P 500) and oil (USO) in Figure 1 where `index_sp` represents S&P 500 price and `index_oil` represents USO price multiplied by 50, the vertical axis is measured in U.S. dollars, and time in the horizontal axis represents the date where time 1 corresponds to 10 May 2007, time 500 to 13 May 2009, time 1000 to 6 May 2011, time 1500 to 3 May 2013, time 2000 to 29 April 2015, and time 2500 to 24 April 2017, respectively.

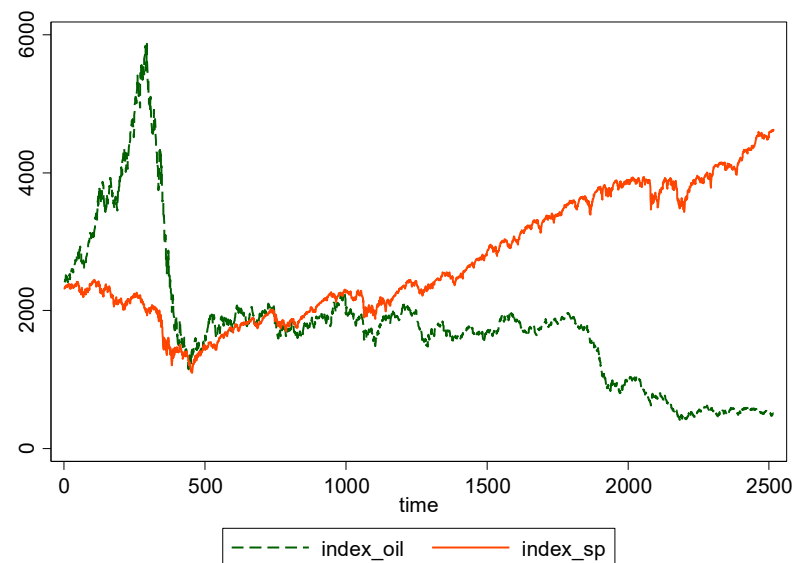


Figure 1. index_sp (S&P 500) and index_oil (USO multiplied by 50). Notes: Daily oil index (USO) multiplied by 50 and stock index (S&P 500) from 10 May 2007 to 16 May 2017 where the vertical axis is measured in U.S. dollars and time in the horizontal axis represents the date where time 1 corresponds to 10 May 2007, time 500 to 13 May 2009, time 1000 to 6 May 2011, time 1500 to 3 May 2013, time 2000 to 29 April 2015, and time 2500 to 24 April 2017, respectively.

Thus, while we use the daily VRP of stock and oil returns between 10 May 2007 and 16 May 2017 to reflect the changes in economic phases, we divide the entire period into five sub-periods and investigate whether and how the VRP of stock and oil are related in each period. The sub-periods are listed in Table 1.

Table 1. Periods and corresponding dates/time.

Whole Period	10 May 2007–16 May 2017 (Time = 1–2516)
Period 1 (Pre-crisis period)	10 May 2007–31 May 2008 (time = 1–266)
Period 2 (Crisis outbreak period)	1 June 2008–30 June 2009 (time = 267–533)
Period 3 (Post-crisis recovery period I)	1 July 2009–31 July 2012 (time = 534–1311)
Period 4 (Post-crisis recovery period II)	1 August 2012–30 September 2014 (time = 1312–1855)
Period 5 (Plunging oil price period)	1 October 2014–16 May 2017 (time = 1856–2516)

We select those periods based partly on Liu et al. (2013), who investigated the dynamic relation among the implied volatilities of stock (VIX), oil (OVX), euro/dollar exchange rate (EVZ), and gold (GVZ) between 3 June 2008 and 20 July 2012. Indeed, Period 2, which is the crisis outbreak period centered on the collapse of Lehman Brothers on 15 September 2008, roughly corresponds to Liu et al.’s (2013) crisis outbreak period, and Period 3, post-crisis recovery period I, corresponds to their post-crisis recovery period so that we can for those periods compare the interaction among implied volatilities shown by Liu et al. (2013) with that of VRP analyzed by this paper. Period 1 is before the outbreak of the global financial crisis. Period 4 is the post-crisis recovery period beyond Liu et al.’s (2013) post-crisis recovery period. Finally, Period 5 is the period of the oil price plunge after the summer of 2014, which may change the relationship between stock and oil VRP.

Figure 2 shows the relationship between VRP_{sp} and VRP_{oil} from 10 May 2007 to 16 May 2017. Here, both stock and oil VRP appear volatile especially during the crisis outbreak

period (Period 2) and are calming in the post-crisis recovery periods (Periods 3 and 4). However, the oil VRP then becomes slightly volatile in accordance with the recent plunging oil prices (in Period 5).

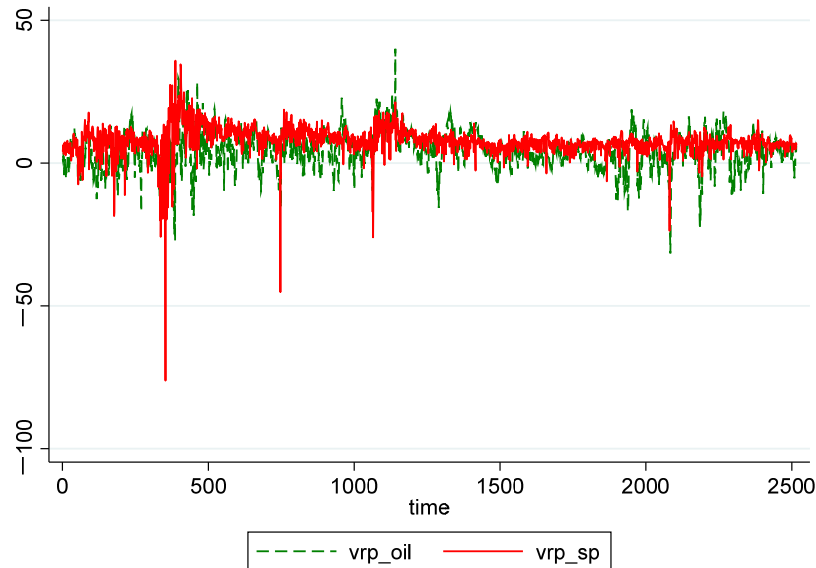


Figure 2. VRP_{oil} and VRP_{sp} from 10 May 2007 to 16 May 2017. Notes: Daily volatility risk premium of oil (VRP_{oil}) and stock (VRP_{sp}) from 10 May 2007 to 16 May 2017. The vertical axis represents their values and time in the horizontal axis represents the date where time 1 corresponds to 10 May 2007, time 500 to 13 May 2009, time 1000 to 6 May 2011, time 1500 to 3 May 2013, time 2000 to 29 April 2015, and time 2500 to 24 April 2017, respectively.

The descriptive statistics of the daily VRP of stock and oil are given in Table 2. It confirms our observation about their volatility as the standard deviation of VRP_{sp} is very high in the crisis outbreak period (10.148 in Period 2), but becomes low after the crisis (4.168, 2.115, and 2.801 in Periods 3, 4, and 5, respectively). The standard deviation of VRP_{oil} is high in the crisis outbreak period (9.754 in Period 2), becomes low in the post-crisis recovering period (4.619 in Period 4), but returns slightly higher in the plunging oil price period (6.724 in Period 5).

Table 2. Descriptive statistics of VRP_{sp} and VRP_{oil} .

	Mean	St. Dev.	Skew.	Kurt.	Corr.	#Obs.
VRP_{sp} (Whole)	7.785	4.891	−3.141	49.116	0.273	2516
VRP_{sp} (Period 1)	6.509	4.794	−1.502	7.724	0.097	266
VRP_{sp} (Period 2)	9.583	10.148	−2.778	22.344	0.278	267
VRP_{sp} (Period 3)	9.536	4.168	−4.291	49.595	0.344	778
VRP_{sp} (Period 4)	6.441	2.115	−0.722	5.156	0.212	544
VRP_{sp} (Period 5)	6.616	2.801	−2.395	24.814	0.222	661
VRP_{oil} (Whole)	4.202	6.627	−0.230	5.047	0.273	2516
VRP_{oil} (Period 1)	2.529	5.842	0.161	2.596	0.097	266
VRP_{oil} (Period 2)	3.623	9.754	−0.061	3.518	0.278	267
VRP_{oil} (Period 3)	5.846	6.317	0.135	4.315	0.344	778
VRP_{oil} (Period 4)	4.231	4.619	0.491	3.134	0.212	544
VRP_{oil} (Period 5)	3.151	6.724	−0.929	5.407	0.222	661

Notes: The upper (resp. lower) part shows descriptive statistics of volatility risk premium of stock, VRP_{sp} , (resp. oil, VRP_{oil}) for the whole period and sub-periods 1, 2, 3, 4, and 5.

The means of VRP_{sp} are between 6.4 and 9.6 and high in the crisis outbreak and just after crisis (9.583 and 9.536 during Periods 2 and 3, respectively), while those of VRP_{oil}

are between 2.5 and 5.8 and highest just after the crisis (5.846 in Period 3). The stock VRP is slightly negatively skewed, whereas that the skewness of the oil VRP can be positive or negative. The kurtosis of the stock VRP is much greater than that of the oil VRP. Finally, the correlation of VRP_{sp} and VRP_{oil} is stable and between 0.2 and 0.3 in all but the first sub-periods.

4. Model Selection

4.1. Unit Root Tests

To select an appropriate model, we begin with unit root tests for VRP_{sp} and VRP_{oil} . The results are presented in Table 3. The null hypothesis of augmented Dickey–Fuller (ADF), Dickey–Fuller–GLS (DF–GLS) and Phillips–Perron (PP) tests is that there is a unit root in the variable.⁸

Table 3. Unit root tests of VRP_{sp} and VRP_{oil} .

	ADF	DF-GLS	PP
VRP_{sp} (Whole)	−11.991 ***	−9.888 ***	−1485.051 ***
VRP_{sp} (Period 1)	−4.928 ***	−4.815 ***	−130.293 ***
VRP_{sp} (Period 2)	−3.623 ***	−3.938 ***	−180.399 ***
VRP_{sp} (Period 3)	−8.022 ***	−7.041 ***	−517.148 ***
VRP_{sp} (Period 4)	−7.744 ***	−7.827 ***	−443.083 ***
VRP_{sp} (Period 5)	−8.864 ***	−8.651 ***	−450.466 ***
VRP_{oil} (Whole)	−11.445 ***	−11.071 ***	−370.718 ***
VRP_{oil} (Period 1)	−4.447 ***	−4.450 ***	−43.848 ***
VRP_{oil} (Period 2)	−4.476 ***	−4.656 ***	−62.622 ***
VRP_{oil} (Period 3)	−5.895 ***	−5.917 ***	−111.345 ***
VRP_{oil} (Period 4)	−3.628 ***	−3.864 ***	−32.445 ***
VRP_{oil} (Period 5)	−5.982 ***	−5.598 ***	−92.555 ***

Notes: *** represents significance at 1% level. All tests reject the null hypothesis at the 1% level significance for all periods.

As Table 3 shows, all tests reject the null hypothesis at the 1% level significance for all periods. Thus, we regard both VRP_{sp} and VRP_{oil} as stationary in the whole and all sub-periods. Note that this is in contrast with the unit root test results on implied volatilities by Liu et al. (2013), where all implied volatilities of stock, oil, gold, and foreign exchange rate have unit roots. Unlike implied volatilities, the VRP of stocks and oil are stationary.

4.2. VAR Model

Since there is no unit root in the whole and all sub-periods, we apply the following VAR model to investigate the dynamic relationship between the VRP of stock and oil.

$$VRP_t = \alpha + \sum_{i=1}^P A_i VRP_t + e_t$$

where $VRP_t = (VRP_{spt}, VRP_{oilt})'$, $\alpha = (\alpha_{sp}, \alpha_{oil})'$, A_i is a 2×2 matrix, P is the lag length, and $e_t = (e_{spt}, e_{oilt})'$ are jointly normally distributed disturbances.

4.3. Choice of Lag Length

We choose the lag length P by comparing the Akaike information criterion (AIC), Hannan and Quinn information criterion (HQIC), and Schwartz’s Bayesian information criterion (SBIC) for each period in the analysis. The results are presented in Table 4.

Table 4. Optimal lag length by AIC, HQIC, and SBIC.

	AIC	HQIC	SBIC	Selected Length
Whole Period	18	5	2	5
Period 1	1	1	1	1
Period 2	2	2	2	2
Period 3	3	3	3	3
Period 4	2	2	2	2
Period 5	7	2	2	2

Notes: The second (resp. third and fourth) column shows the optimal lag length given by AIC (resp. HQIC and SBIC) for each period/sub-period. The fifth column shows the lag length that is used in the analysis of this paper.

For Periods 1, 2, 3, and 4, all AIC, HQIC, and SBIC criteria have the same results, which we choose as the lag length in the analysis. In contrast, for the entire period and Period 5, the optimal lag lengths given by the different criteria do not match. In particular, AIC tends to provide a larger optimal lag length. Nonetheless, because the values of the AIC (resp. SBIC) for lag lengths 5 and 18 (resp. 5 and 2) are close and the lag length 5 given by HQIC is in the middle of the three criteria, we select the lag length as 5 for the entire period. Likewise, because the AIC values for lag lengths 2 and 7 are very close and HQIC and SBIC give the same length of 2, we set the lag length to 2 in the analysis of Period 5.⁹ Consequently, we select the optimal lag lengths given by the HQIC for all periods in this study.

5. Empirical Results

5.1. Results for the Whole Period

Table 5 reports the results of the Granger causality tests for the entire period. Both test statistics are significant at the 1% level. Thus, for the entire period, the VRP of stock and oil dynamically influence each other in the sense of Granger causality.

Table 5. Granger causality test (whole period).

Null Hypothesis	Period	Chi 2	# of Lags
VRP_{sp} does not GC VRP_{oil}	Whole	21.174 ***	5
VRP_{oil} does not GC VRP_{sp}	Whole	26.076 ***	5

Notes: VRP_{oil} (resp. VRP_{sp}) represents volatility risk premium of oil (resp. stock). *** indicates significance at 1% level.

Figure 3 shows the orthogonalized impulse response functions of stock and oil VRP with 95% confidence intervals where we order VRP_{oil} before VRP_{sp} .¹⁰ The impulse response functions of VRP_{sp} to VRP_{oil} shown in the above-right graph are significantly positive until date 14, as the lower bounds of their 95% confidence intervals are larger than zero, and gradually decrease toward date 20. Likewise, the impulse response functions of VRP_{oil} to VRP_{sp} shown in the below-left graph are significantly positive until date 17, except for dates 3 and 4, and tend to decrease toward date 20. Thus, if we look at the whole period, shocks in both VRP_{sp} and VRP_{oil} have, though small, significantly positive effects on each other for most of the 20 trading days (about 1 month) after the shock.

The variance decomposition results are listed in Table 6. Shocks to VRP_{sp} explained by innovations in VRP_{oil} are shown in the fourth column, which indicates that 5.2% of the forecast-error variance of VRP_{sp} is explained by innovations in VRP_{oil} on date 20. Meanwhile, shocks to VRP_{oil} explained by innovations in VRP_{sp} are shown in the third column, which indicates that 2.7% of the forecast-error variance of VRP_{oil} is explained by innovations in VRP_{sp} on date 20.

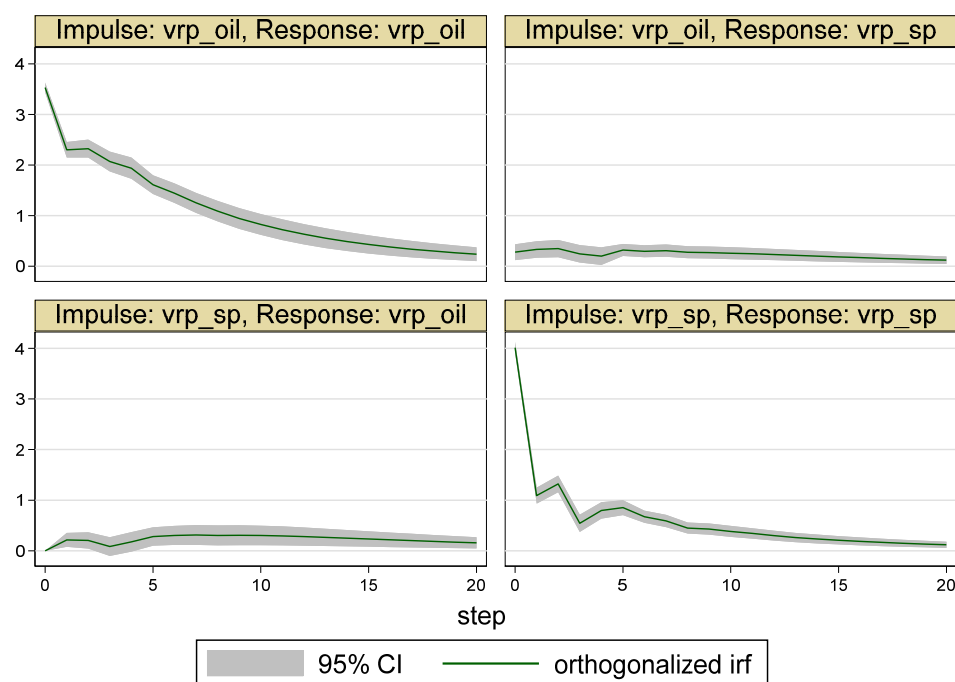


Figure 3. Impulse response functions (whole period). Notes: Vrp_oil (resp. vrp_sp) represents volatility risk premium of oil (resp. stock). The solid line represents orthogonalized impulse response functions and the gray area represents their 95% confidence intervals. The above-right graph shows the impulse response functions of vrp_sp to vrp_oil. The below-left graph shows the impulse response functions of vrp_oil to vrp_sp.

Table 6. Variance decompositions (whole period).

Impulse	VRP _{oil}	VRP _{sp}	VRP _{oil}	VRP _{sp}
Response	VRP _{oil}		VRP _{sp}	
1	1	0	0.005	0.995
5	0.996	0.004	0.020	0.980
10	0.986	0.014	0.037	0.963
15	0.977	0.023	0.048	0.952
20	0.973	0.027	0.052	0.948

Note: The first column shows the length of horizon. The third (resp. second) column shows the variance decomposition of VRP_{oil} to VRP_{sp} (resp. itself). Also, the fourth (resp. fifth) column shows the variance decomposition of VRP_{sp} to VRP_{oil} (resp. itself).

Thus, for the whole period, the VRP of stock and oil have similar (i.e., small but significantly positive) effects on each other. That is, investors’ sentiments in the stock and oil markets affect each other similarly in the sense that an increase in the premium for volatility risk in one market propagates to the other market, although the effect is not large.

5.2. Results for the Sub-Periods

If we see the relations between the VRP of stock and oil in the sub-periods, however, we have rather different pictures. Table 7 shows the results of the Granger causality tests for the sub-periods.

It is interesting that the stock VRP Granger causes the oil VRP very strongly at the 1% significance level in the post-crisis recovery period I (i.e., Period 3), but not in the other sub-periods except for the plunging oil price period (i.e., Period 5) in which VRP_{sp} Granger causes VRP_{oil} only at 10% significance level. In contrast, the oil VRP Granger causes the stock VRP strongly at the 1% significance level in the post-crisis recovery periods (i.e., periods 3 and 4) and relatively strongly at the 5% significance level in the crisis outbreak and the plunging oil price periods (i.e., Periods 2 and 5, respectively). Thus, while the

Granger causality from VRP_{sp} to VRP_{oil} over the entire period is mainly from that in post-crisis recovery period I (i.e., Period 3), the Granger causality from VRP_{oil} to VRP_{sp} is persistent after the outbreak of the crisis.

Table 7. Granger causality tests for sub-periods.

Null Hypothesis	Period	Chi 2	# of Lags
VRP_{sp} does not GC VRP_{oil}	Period 1	1.214	1
	Period 2	1.011	2
	Period 3	35.073 ***	3
	Period 4	1.439	2
	Period 5	4.786 *	2
VRP_{oil} does not GC VRP_{sp}	Period 1	0.024	1
	Period 2	6.861 **	2
	Period 3	27.029 ***	3
	Period 4	18.452 ***	2
	Period 5	9.066 **	2

Notes: VRP_{oil} (resp. VRP_{sp}) represents volatility risk premium of oil (resp. stock). ***, **, and * indicate significance at 1%, 5%, and 10% levels, respectively.

The results of the Granger causality test suggest that the Granger causality is stronger and more persistent from oil to stock than from stock to oil. We show only the results for which VRP_{oil} is ordered before VRP_{sp} in the following analyses of orthogonalized impulse response functions and variance decomposition.

The impulse response functions for each sub-period are shown in Figures 4–8. As the graphs at the bottom-left of Figures 4–8 show, in the sub-periods the stock VRP has little effect on the oil VRP. Indeed, there are no significant effects from VRP_{sp} to VRP_{oil} except in the post-crisis recovery period I (i.e., Period 3), but even in that period the effects are short-lived and significant only up to the 2nd trading day after the shock, as the lower-bounds of the 95% confidence level of the below-left graph in Figure 6 show.

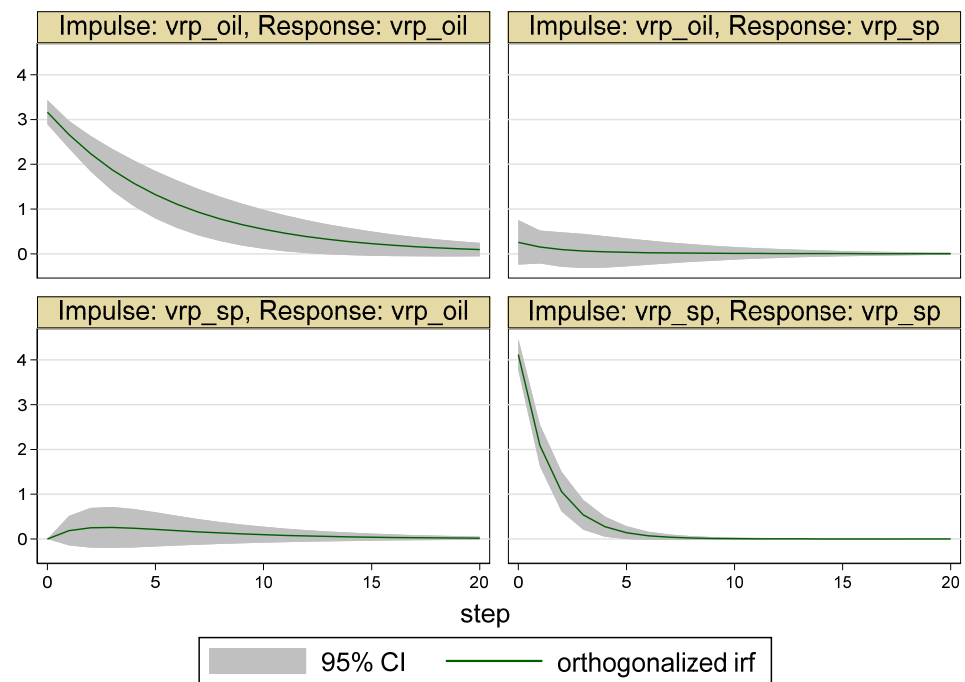


Figure 4. Impulse Response Function (Period 1).

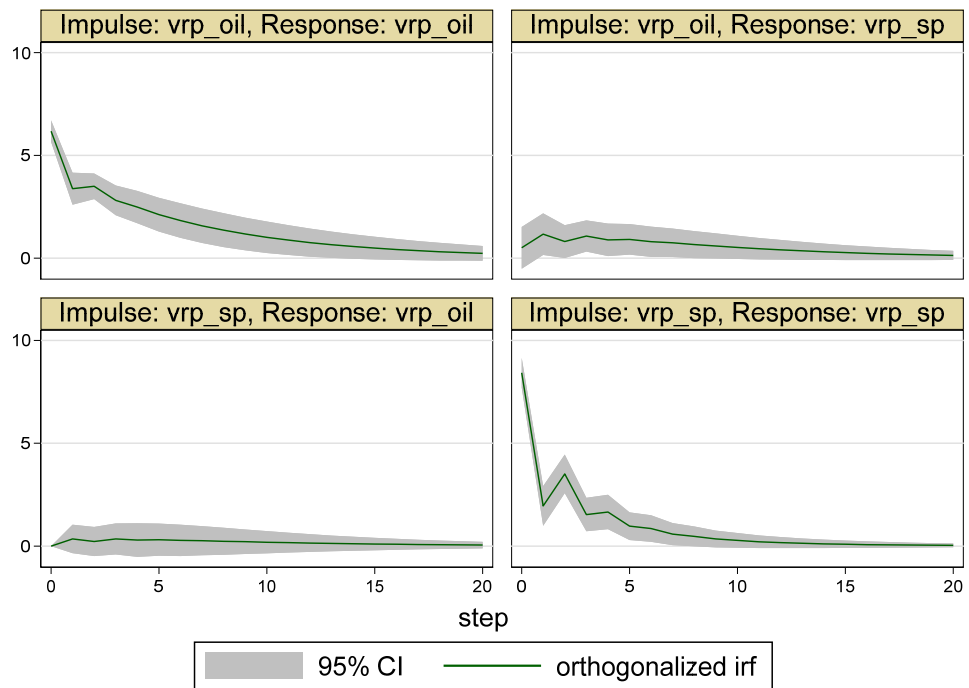


Figure 5. Impulse Response Function (Period 2).

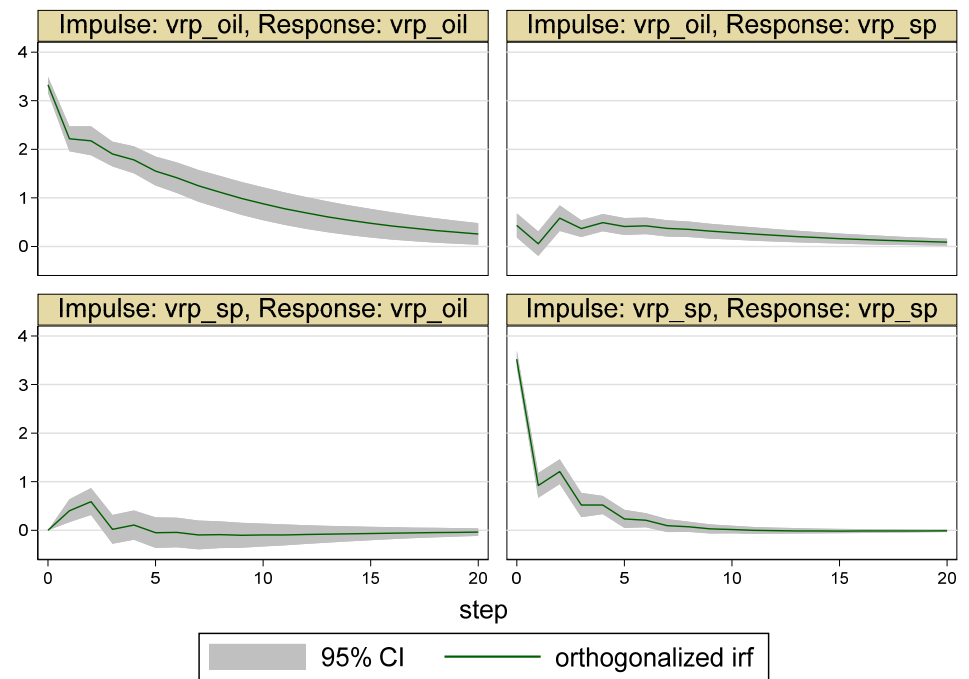


Figure 6. Impulse Response Function (Period 3).

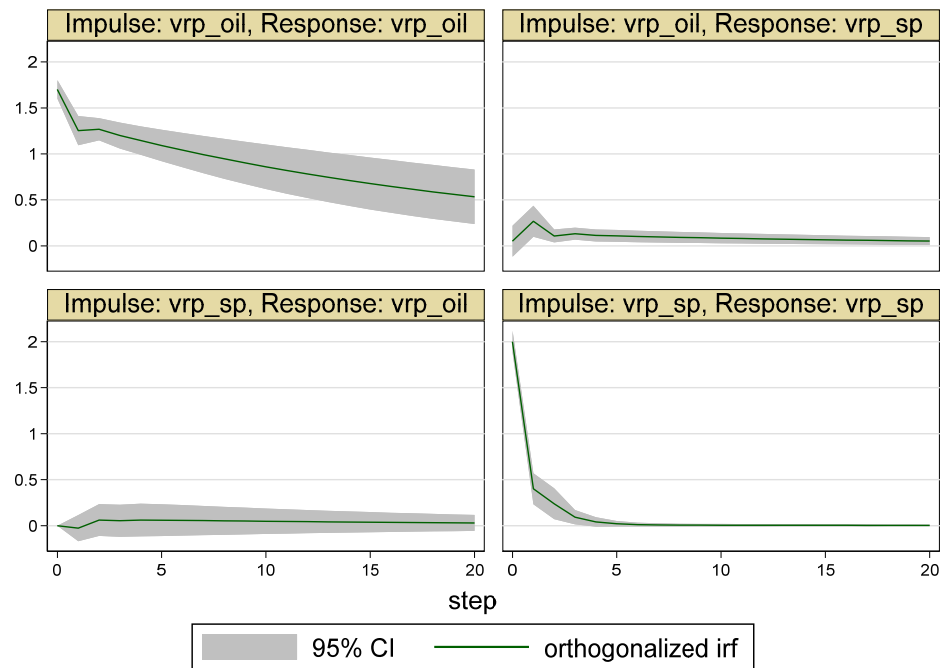


Figure 7. Impulse Response Function (Period 4).

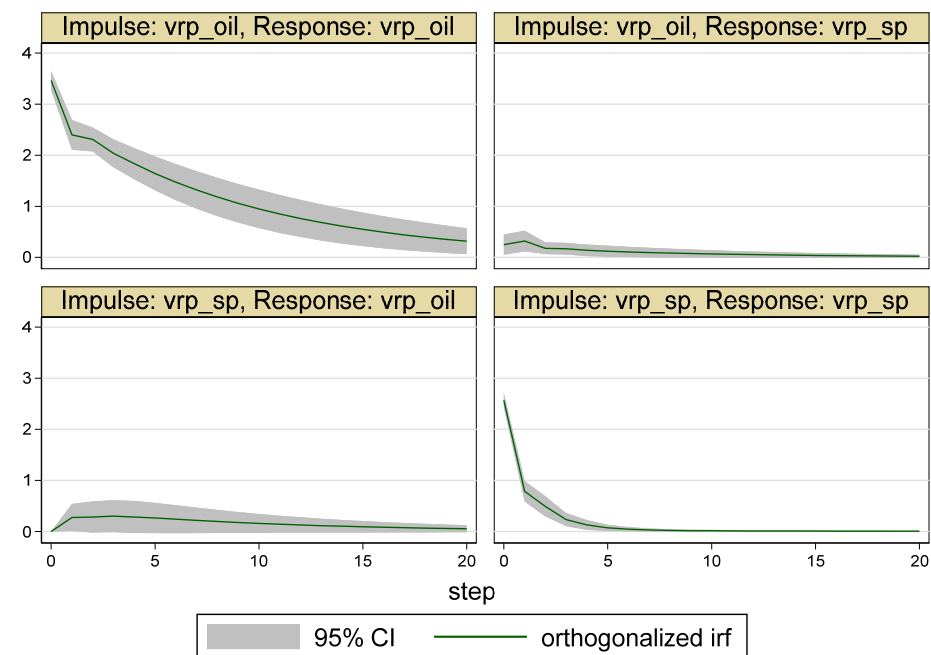


Figure 8. Impulse Response Function (Period 5). Notes: Vrp_{oil} (resp. vrp_{sp}) represents volatility risk premium of oil (resp. stock). The solid line represents orthogonalized impulse response functions and the gray area represents their 95% confidence intervals. The above-right graph shows the impulse response functions of vrp_{sp} to vrp_{oil} . The below-left graph shows the impulse response functions of vrp_{oil} to vrp_{sp} .

In contrast, as the graphs on the above-right of Figures 5–8 show, the oil VRP has, though small, significantly positive and long-lasting effects on stock VRP after the outbreak of the crisis (i.e., in Periods 2, 3, 4, and 5). For example, in the crisis outbreak period (i.e., period 2), as the above-right graph of Figure 5 shows, the orthogonalized impulse response functions from VRP_{oil} to VRP_{sp} are significantly positive from the 1st to the 8th trading days after the shock. In post-crisis recovery period I (i.e., Period 3), as shown in Figure 6, they are significantly positive in all but the 1st trading day. In post-crisis recovery period II

(Period 4), as shown in Figure 7, they are significantly positive in all trading days after the shock. Finally, in the plunging oil price period (i.e., Period 5), as shown in Figure 8, they are significantly positive up to the 7th trading day. Consistent with the Granger causality tests, those results of the impulse response functions in the sub-periods suggest that the oil VRP dynamically affects the stock VRP, but not vice versa.

Table 8 shows the results of the variance decomposition for the sub-periods. The shocks to VRP_{oil} explained by innovations in VRP_{sp} are shown in the third column, and the shocks to VRP_{sp} explained by innovations in VRP_{oil} are shown in the fourth column. Similar to the results of the Granger causality tests and impulse response functions, the forecast error variances of the oil VRP explained by innovations in the stock VRP are much smaller than those of the stock VRP explained by the oil VRP in all sub-periods except the pre-crisis period (Period 1). For example, on date 20, the forecast-error variances of VRP_{oil} (VRP_{sp}) explained by innovations in VRP_{sp} (VRP_{oil}) are 0.011 (0.005), 0.009 (0.080), 0.015 (0.114), 0.002 (0.048), and 0.015 (0.040) in periods 1, 2, 3, 4, and 5, respectively. Again, the oil VRP dynamically affects the stock VRP much more than the stock VRP does, which dynamically affects the oil VRP after the financial crisis.

Table 8. Variance decompositions (sub-periods).

Impulse Response	VRP_{oil}	VRP_{sp}	VRP_{oil}	VRP_{sp}
	VRP_{oil}		VRP_{sp}	
Period 1				
1	1	0	0.004	0.996
5	0.992	0.008	0.005	0.995
10	0.989	0.011	0.005	0.995
15	0.989	0.011	0.005	0.995
20	0.989	0.011	0.005	0.995
Period 2				
1	1	0	0.004	0.996
5	0.995	0.005	0.044	0.956
10	0.992	0.008	0.070	0.930
15	0.991	0.009	0.078	0.922
20	0.991	0.009	0.080	0.920
Period 3				
1	1	0	0.015	0.985113
5	0.981	0.019	0.056	0.944
10	0.985	0.015	0.095	0.905
15	0.985	0.015	0.110	0.890
20	0.985	0.015	0.114	0.886
Period 4				
1	1	0	0.001	0.999
5	0.999	0.001	0.027	0.973
10	0.998	0.002	0.037	0.963
15	0.998	0.002	0.044	0.956
20	0.998	0.002	0.048	0.952
Period 5				
1	1	0	0.009	0.991
5	0.990	0.010	0.032	0.968
10	0.986	0.014	0.037	0.963
15	0.985	0.015	0.039	0.961
20	0.985	0.015	0.040	0.960

Note: VRP_{oil} (resp. VRP_{sp}) represents volatility risk premium of oil (resp. stock). The first column shows the length of horizon. The third (resp. second) column shows the variance decomposition of VRP_{oil} to VRP_{sp} (resp. itself). Also, the fourth (resp. fifth) column shows the variance decomposition of VRP_{sp} to VRP_{oil} (resp. itself).

Thus, the analyses of the sub-periods reveal that the dynamic relationship between the stock and oil VRP depends on the economic situation and, more importantly, that the VRP of oil has long-lasting and significantly positive effects on that of stock after the outbreak of the financial crisis, whereas the effects of the stock VRP on the oil VRP are limited and, if any, much more short-lived. That is, investors' sentiments propagate from the oil market to

the stock market after the global financial crisis, but not from the stock market to the oil market, except in the first half of the recovery period after the financial crisis.

6. Robustness Analysis

To check the robustness of the results in orthogonalized impulse response functions and variance decomposition, we repeat the analysis by reversing the order of the VRP to place VRP_{sp} before VRP_{oil} . The results are quite similar to those above, although the effects of VRP_{sp} on VRP_{oil} (resp. VRP_{oil} on VRP_{sp}) become slightly stronger (resp. weaker) than those previously obtained.

For example, Figures 9–13 show the orthogonalized impulse response functions for the sub-periods. The figures are qualitatively the same as those above.

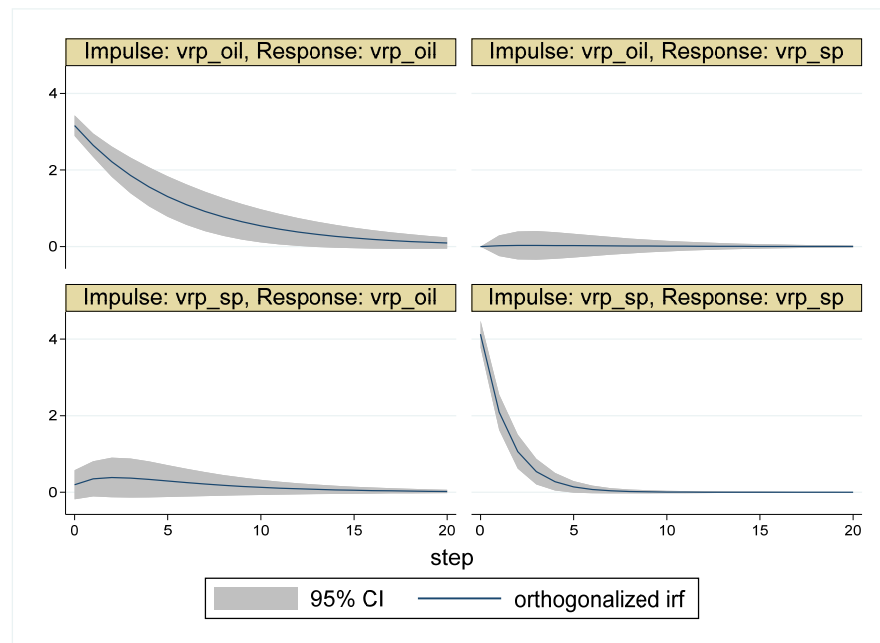


Figure 9. Impulse Response Function with Order VRP_{sp} before VRP_{oil} (Period 1).

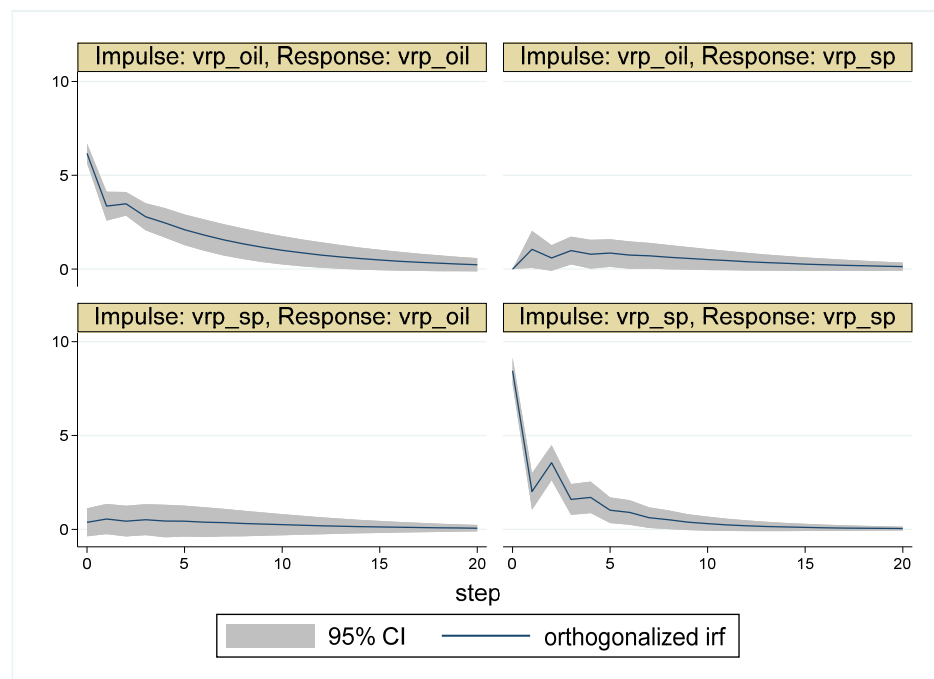


Figure 10. Impulse Response Function with Order VRP_{sp} before VRP_{oil} (Period 2).

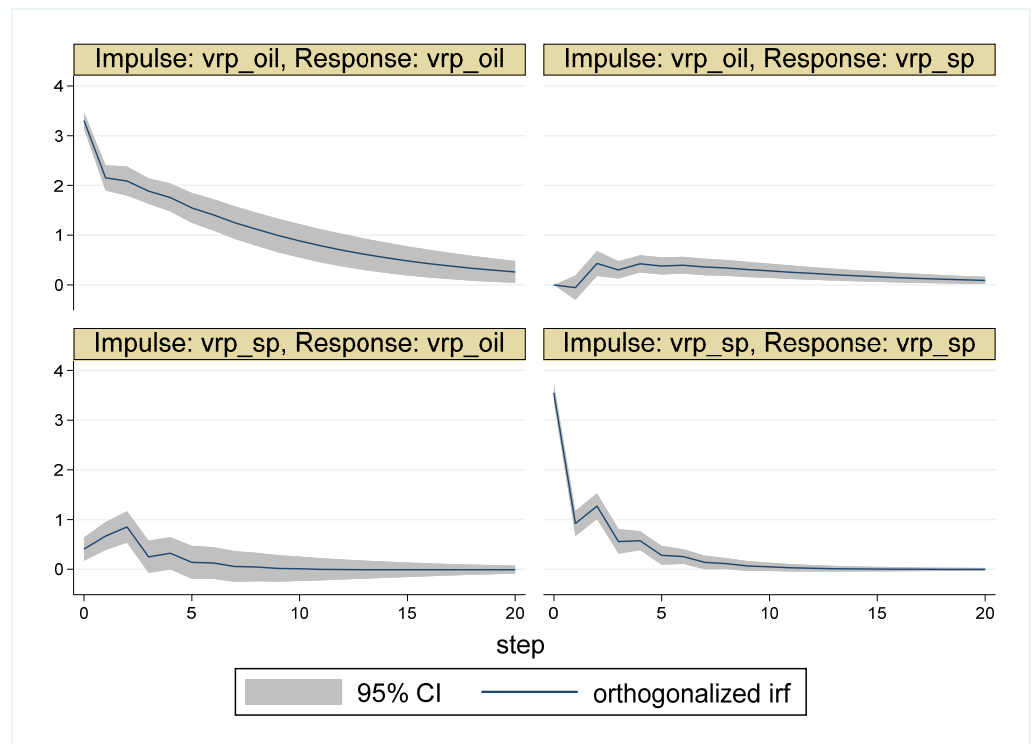


Figure 11. Impulse Response Function with Order VRP_{sp} before VRP_{oil} (Period 3).

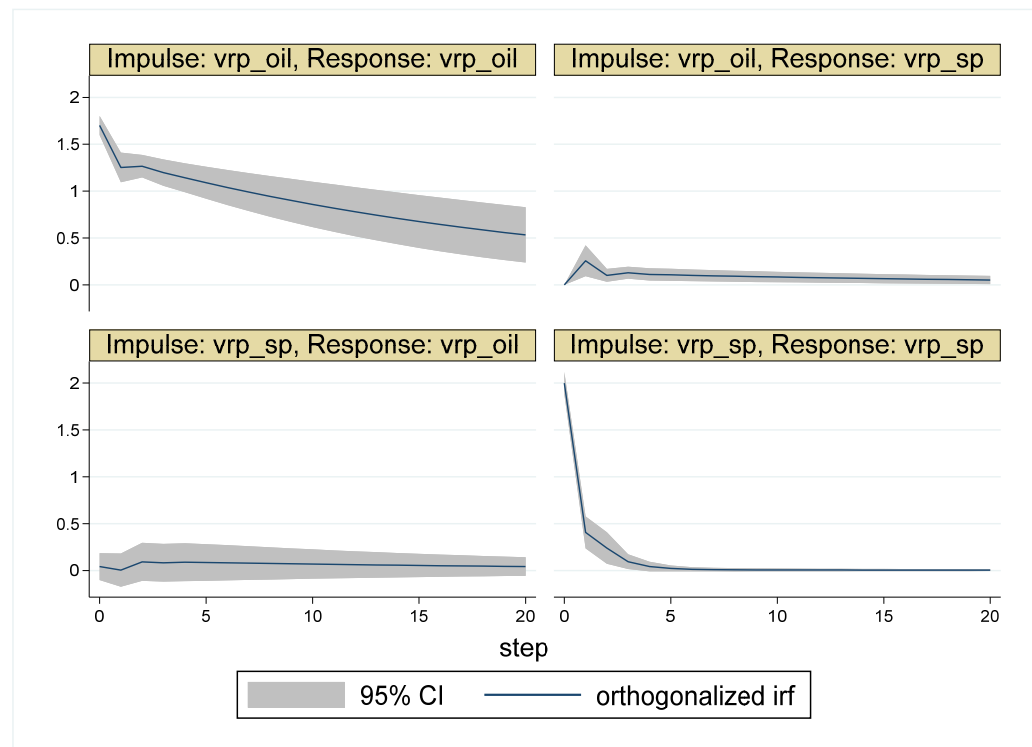


Figure 12. Impulse Response Function with Order VRP_{sp} before VRP_{oil} (Period 4).

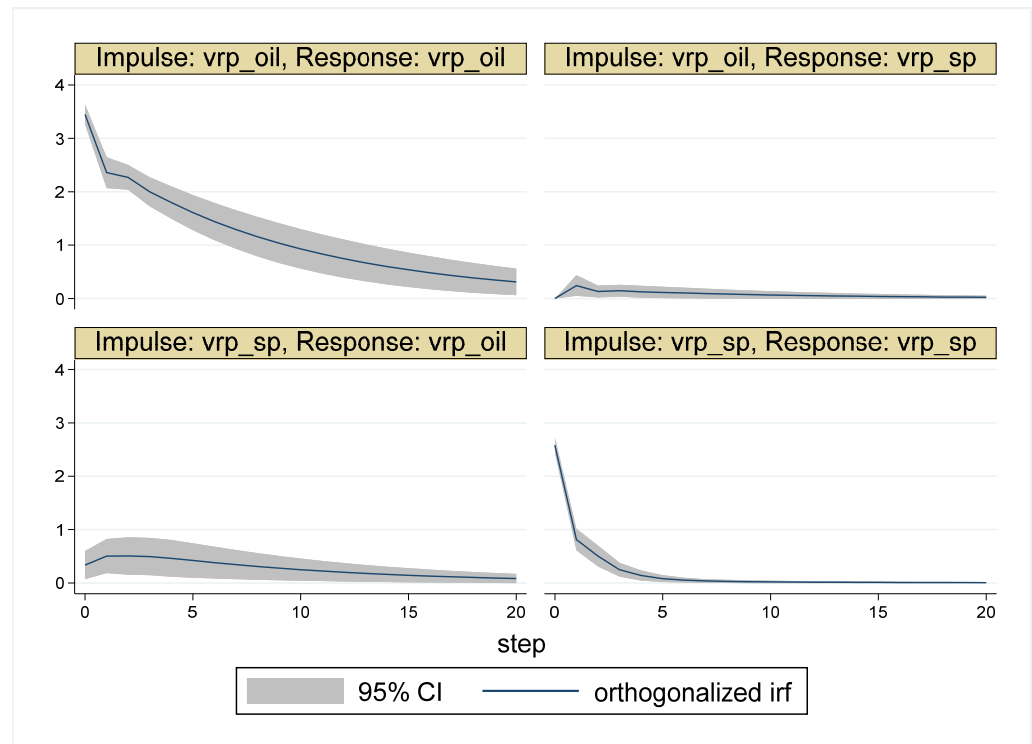


Figure 13. Impulse Response Function with Order VRP_{sp} before VRP_{oil} (Period 5). Notes: Vrp_{oil} (resp. vrp_{sp}) represents volatility risk premium of oil (resp. stock). The solid line represents orthogonalized impulse response functions and the gray area represents their 95% confidence interval. The above-right graph shows the impulse response functions of vrp_{sp} to vrp_{oil} . The below-left graph shows the impulse response functions of vrp_{oil} to vrp_{sp} .

Table 9 shows the results of the variance decomposition focusing on the 20th trading day after the shocks. This table shows that the forecast-error variances of oil (resp. stock) VRP explained by the innovation in the stock (resp. oil) VRP are now generally larger (resp. smaller) than those in the previous case. For example, in Period 3, the forecast error variance of VRP_{oil} (resp. VRP_{sp}) explained by VRP_{sp} (resp. VRP_{oil}) is 0.039 (resp. 0.084) on the 20th day after the shock, whereas the corresponding values in the case above are 0.015 (resp. 0.114).

Table 9. Variance decompositions with order VRP_{sp} before VRP_{oil} . (Effects on the 20th trading day after the shock).

Impulse	VRP_{oil}	VRP_{sp}	VRP_{oil}	VRP_{sp}
Response	VRP_{oil}		VRP_{sp}	
Period 1 20	0.975	0.025	0	1
Period 2 20	0.979	0.021	0.065	0.935
Period 3 20	0.961	0.039	0.084	0.916
Period 4 20	0.995	0.005	0.045	0.955
Period 5 20	0.956	0.044	0.023	0.977

Note: VRP_{oil} (resp. VRP_{sp}) represents volatility risk premium of oil (resp. stock). The first column shows the length of horizon. The third (resp. second) column shows the variance decomposition of VRP_{oil} to VRP_{sp} (resp. itself). Also, the fourth (resp. fifth) column shows the variance decomposition of VRP_{sp} to VRP_{oil} (resp. itself).

Similarly, we repeat the analysis using different lag lengths, although not optimal, to see how the results can change. We choose lag lengths equal to 4 and 9 because the former corresponds roughly to 1 week (5 trading days) coverage and the latter to 2 weeks (10 trading days). We do not report the results here, but they are qualitatively the same as those obtained above while the effects of stock on oil seem to be slightly stronger.

7. Summary and Concluding Remarks

The VRP was found to have predictive power for returns in many different assets. While most extant studies have analyzed the predictability of VRP on asset returns, this study investigated how the VRP of different assets, specifically those of stock and oil, are dynamically related to each other. To this end, we obtained stock VRP as the difference between the VIX published by the CBOE and the realized volatility of the S&P 500 stock index provided by the Oxford–Man Institute of Quantitative Finance. In contrast, to construct the oil VRP, we estimated the realized volatility of the USO by a stochastic volatility model and subtracted it from the OVX published by the CBOE.

Using daily data from 10 May 2007 to 16 May 2017, we conducted VAR analyses on the stock and oil VRP for the whole period and five sub-periods that represent the pre-crisis, crisis outbreak, post-crisis recovery (the first and the second half) and plunging oil price periods.

The analysis of the whole period shows that the VRP of stock and oil have similar (i.e., small but significantly positive) effects on each other. However, the analyses of the five sub-periods revealed a different picture. The dynamic relationship between the stock and oil VRP depends on the economic situation and, contrary to the results for the whole period, the effects of the stock and oil VRP on their counterparts are quite different: The effects of the stock VRP on the oil VRP are limited mainly in the first half of the post-crisis recovery period and are short-lived. In contrast, the VRP of oil has significantly positive and long-lasting effects on that of stock in all sub-periods after the outbreak of the financial crisis.

It is worth pointing out that those results suggest that the investors' sentiments (measured by volatility risk premia) are transmitted from the oil market to the stock market over time, but not the other way around. While Christoffersen and Pan (2018) find the predictability of oil implied volatility on stock returns and implied volatility, the relationship between oil and stock VRP is still an unexplored point in the extant literature and is a rather unexpected finding because the financialization of commodities means a massive increase in investment in commodities by the investors in the traditional stock and bond markets; hence, the direction of the effects is thought to be from the stock market to the commodity market, and not from the commodity market to the stock market.

However, the mechanism of such a transmission of VRP from oil to stock has not yet been elucidated. One possible channel is the funding constraints of financial intermediaries. Christoffersen and Pan (2018) found that increases in oil implied volatility predict tightening funding constraints of financial intermediaries, which can affect stock price and implied volatility; the oil VRP may affect the stock VRP through institutional investors' funding constraints. Hattori et al. (2021) found that increases in the U.S. stock VRP tend to reduce the fund flow into stocks of emerging economy countries and suggested that this can be the cause of the spillover of VRP among countries. One may also speculate that the tendency of declining oil prices after the financial crisis and the emergence of shale oil/gas may make investments in U.S. stocks by petroleum-exporting countries (i.e., "petrodollars") more sensitive to the future uncertainty of oil prices, to which other investors become more sensitive. Investigating what causes oil VRP to affect stock VRP is an important target for future research.

In addition, owing to the constraint on data availability, we must estimate the realized volatility by applying a stochastic volatility model to the daily data of the USO returns. This makes our estimation of the oil VRP, which is the difference between the OVX published by the CBOE and the realized volatility calculated, prone to the misspecification of the

used model. To utilize the same estimation method as the stock VRP, or more specifically as the stock-realized volatility, it is desirable to estimate the realized volatility of the USO using intraday 5 min return data. Such a method also enables us to investigate the relationship among the returns and VRP of stock and oil. This is another important topic for future research.

Finally, analyzing the relationship among the returns and VRP of stock and oil helps investors and policy makers understand how the returns and sentiment of one market affect those of the other market. Indeed, the results of this study show that, after the global financial crisis, the shock in investor sentiment on oil prices propagates to that on stock prices, but not vice versa. This means that it is important for investors and policy makers to pay more attention to the spillover of shocks from oil to stock than from stock to oil in order to attain better risk management and asset allocation. Therefore, the direction of this study is fruitful.

Author Contributions: Conceptualization, N.N. and K.O.; Methodology, N.N., K.O. and D.Y.; Formal analysis, N.N., K.O. and D.Y.; Data curation, D.Y.; Writing—original draft, K.O.; Writing—review & editing, K.O.; Project administration, K.O.; Funding acquisition, N.N. and K.O. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Japan Society for the Promotion of Science grant number 21H00727.

Data Availability Statement: We calculate the daily stock VRP as the difference between the VIX index and the daily realized volatility of the S&P 500 stock index. The data of VIX index can be obtained from the CBOE home page <https://www.cboe.com/us/indices/dashboard/VIX/> (accessed on 26 December 2022). The data of daily realized volatility of the S&P 500 stock index is provided by the Oxford-Man Institute of Quantitative Finance <http://realized.oxford-man.ox.ac.uk/> (accessed on 26 December 2022). The daily oil VRP is calculated as the difference between the OVX and the daily realized volatility of USO, which is the underlying asset of OVX. The data of OVX is available from the CBOE home page <https://www.cboe.com/us/indices/dashboard/OVX/> (accessed on 26 December 2022). The daily realized volatility is estimated by the stochastic volatility model (explained in Appendix A) using the data of daily USO returns, which is purchased from Refinitiv.

Acknowledgments: This study was conducted in the project “Economic and Financial Analysis of Commodity Markets” at the Research Institute of Economy, Trade and Industry (RIETI). We are grateful to Katja Ignatieva, Kentaro Iwatsubo, Katsushi Nakajima, Hiroshi Ohashi, Tatsuyoshi Okimoto, Makoto Takahashi, Tianyang Wang, Yuji Yamada, Makoto Yano, and all seminar participants at the AJRC and RIETI workshops on economic and financial analysis of commodity markets in ANU, the Discussion Paper seminar at RIETI (RIETI DP 18-E-027), the fourth annual J.P. Morgan Center for Commodities International Symposium, and anonymous referees for their helpful comments and discussions. Nakamura and Ohashi acknowledge financial support from Grants-in-Aid for Scientific Research (21H00727) by Japan Society for the Promotion of Science.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A. Estimation of Realized Volatility by a Stochastic Volatility Model

We estimate the daily realized volatility of oil by applying the following stochastic volatility model to its return.

$$\begin{aligned}
 r_{oil\ t} &= \mu_{oil} + \beta_{oil} \left(e^{\frac{h_t}{2}} - OVX_t \right) + e^{\frac{h_t}{2}} \epsilon_t, \\
 h_t &= \mu_h + \beta_h (h_{t-1} - \mu_h) + \sigma_h \eta_t, \\
 \epsilon_t &\equiv \sqrt{\frac{\nu-2}{\nu}} \frac{\xi_t}{\sqrt{\zeta_t}}, \quad \zeta_t \sim \Gamma(\nu/2, \nu/2), \\
 \begin{pmatrix} \xi_t \\ \eta_t \end{pmatrix} &\sim N \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix} \right)
 \end{aligned}$$

where $r_{oil\ t}$ denotes the daily return of the USO, $N()$ denotes the bivariate standard normal distribution with correlation ρ , and $\Gamma(\nu/2, \nu/2)$ denotes the gamma-distribution

with shape and scale parameters equal to $\nu/2$. Here, realized volatility is estimated as $RV_{oil\ t} = e^{\frac{h_t}{2}}$.

We estimated the model for the five sub-periods using Bayesian statistical inference according to the widely applicable information criterion (WAIC). The estimates of mean, 5%, and 95% percentile points (in parentheses) are as follows:

Table A1. Estimates of mean and standard deviation of parameters.

Parameter	Period 1	Period 2	Period 3	Period 4	Period 5
μ_{oil}	0.305 (0.192, 0.415)	-0.336 (-0.496, -0.178)	-0.017 (-0.03, -0.01)	0.007 (0.003, 0.010)	-0.194 (-0.270, -0.122)
β_{oil}	0.041 (-0.039, 0.123)	0.068 (-0.017, 0.151)	0.032 (-0.044, 0.109)	0.028 (-0.046, 0.101)	0.085 (0.004, 0.164)
μ_h	1.232 (0.936, 1.530)	2.472 (2.084, 2.820)	1.278 (1.134, 1.428)	0.292 (0.095, 0.509)	1.675 (1.413, 1.939)
β_h	0.849 (0.774, 0.915)	0.891 (0.835, 0.940)	0.809 (0.733, 0.873)	0.846 (0.771, 0.903)	0.914 (0.876, 0.946)
σ_h	0.380 (0.336, 0.431)	0.369 (0.325, 0.417)	0.363 (0.325, 0.406)	0.368 (0.328, 0.413)	0.349 (0.311, 0.391)
ρ	-0.488 (-0.630, -0.335)	-0.4649 (-0.608, -0.310)	-0.467 (-0.587, -0.331)	-0.457 (-0.595, -0.308)	-0.413 (-0.550, -0.269)
ν	8.047 (6.455, 9.753)	8.083 (6.520, 9.757)	8.455 (6.972, 10.116)	8.014 (6.476, 9.664)	8.534 (6.994, 10.210)

Notes

- 1 Results by extant researches about the effects of oil volatility-related variables on stock returns and volatility are mixed. For example, Ornelas and Mauad (2019) find little predictability of oil VRP on S&P 500 returns, Bams et al. (2017) find that difference of oil VRP is priced only on returns of oil-related stocks, and Christoffersen and Pan (2018) find predictability of oil implied volatility on stock returns and implied volatility.
- 2 Volatility swap and variance swap, where variance is the square of volatility, are traded in over-the-counter derivative markets.
- 3 Ornelas and Mauad (2019) explain what kind of realized volatility is used in the literature to approximate the expected future volatility.
- 4 <https://www.cboe.com/us/indices/dashboard/VIX/> (3 March 2022).
- 5 <http://realized.oxford-man.ox.ac.uk/> (3 March 2023).
- 6 <https://www.cboe.com/us/indices/dashboard/OVX/> (3 March 2023).
- 7 Appendix A explains how we estimate the realized volatility of oil.
- 8 For Augmented Dickey–Fuller (ADF), Dickey–Fuller–GLS (DF–GLS), and Phillips–Perron (PP) tests, see Dickey and Fuller (1979); Elliott et al. (1996); and Phillips and Perron (1988), respectively.
- 9 Analyses with different lag length provide results quite similar to those in this paper.
- 10 We obtain the similar result if we reverse the order of VRP_{oil} and VRP_{sp} . We select this ordering since the results of Granger causality tests show more persistent Granger causality from oil to stock than from stock to oil for most of the sub-periods. For more detail, see next subsection.

References

Bams, Dennis, Gildas Blanchard, Iman Honarvar, and Thorsten Lehnert. 2017. Does oil and gold price uncertainty matter for the stock market? *Journal of Empirical Finance* 44: 270–85. [CrossRef]

Bollerslev, Tim, George Tauchen, and Hao Zhou. 2009a. Expected Stock Returns and Variance Risk Premia. *Review of Financial Studies* 22: 4463–92. [CrossRef]

Bollerslev, Tim, James Marrone, Lai Xu, and Hao Zhou. 2009b. Stock Return Predictability and Variance Risk Premia: Statistical Inference and International Evidence. *Journal of Financial and Quantitative Analysis* 60: 633–61. [CrossRef]

Bouri, Elie, Brian Lucey, and David Roubaud. 2020. Dynamic and determinants of spillovers across the option-implied volatilities of US equities. *Quarterly Review of Economics and Finance* 75: 257–64. [CrossRef]

- Bouri, Elie, Xiaojie Lei, Yahua Xu, and Hongwei Zhang. 2023. Connectedness in implied higher-order moments of precious metals and energy markets. *Energy* 263: 125588. [CrossRef]
- Christoffersen, Peter, and Xuhui (Nick) Pan. 2018. Oil volatility risk and expected stock returns. *Journal of Banking and Finance* 95: 5–26. [CrossRef]
- Della Corte, Pasquale, Tarun Ramadorai, and Lucio Sarno. 2016. Volatility risk premia and exchange rate predictability. *Journal of Financial Economics* 120: 21–40. [CrossRef]
- Dickey, David A., and Wayne A. Fuller. 1979. Distribution of the Estimators for Autoregressive Time Series with a Unit Root. *Journal of the American Statistical Association* 74: 427–31. [CrossRef]
- Diebold, Francis X., and Kamil Yilmaz. 2014. On the network topology of variance decompositions: Measuring the connectedness of financial firms. *Journal of Econometrics* 182: 119–34. [CrossRef]
- Dutta, Anupam, Elie Bouri, and David Roubaud. 2019. Nonlinear relationships amongst the implied volatilities of crude oil and precious metals. *Resources Policy* 61: 473–78. [CrossRef]
- Elliott, Graham, James H. Stock, and Thomas J. Rotemberg. 1996. Efficient Tests for an Autoregressive Unit Root. *Econometrica* 64: 813–36. [CrossRef]
- Gagnon, Marie-Helene, Gabriel J. Power, and Dominique Toupin. 2015. Dynamics between crude oil and equity markets under the risk-neutral measure. *Applied Economics Letters* 22: 370–77. [CrossRef]
- Hattori, Masazumi, Ilhoyock Shim, and Yoshihiko Sugihara. 2021. Cross-stock market spillovers through variance risk premiums and equity flows. *Journal of International Money and Finance* 119: 102480. [CrossRef]
- Iqbal, Najaf, Elie Bouri, Guangrui Liu, and Ashish Kumar. 2022. Volatility spillovers during normal and high volatility states and their driving factors: A cross-country and cross-asset analysis. *International Journal of Finance and Economics*, 1–21. [CrossRef]
- Liu, Ming-Lei, Qiang Ji, and Ying Fan. 2013. How does all market uncertainty interact with other markets? An empirical analysis of implied volatility index. *Energy* 55: 860–68. [CrossRef]
- Londono, Juan, and Hao Zhou. 2016. Variance Risk Premiums and the Forward Premium Puzzle. *Journal of Financial Economics* 124: 415–40. [CrossRef]
- Ohashi, Kazuhiko, and Tatsuyoshi Okimoto. 2016. Increasing Trends in the Excess Comovement of Commodity Prices. *Journal of Commodity Markets* 1: 48–64. [CrossRef]
- Ornelas, Jose Renato Haas, and Roberto Baltieri Mauad. 2019. Volatility risk premia and futures commodities returns. *Journal of International Money and Finance* 96: 341–60. [CrossRef]
- Phillips, Peter C. B., and Pierre Perron. 1988. Testing for a Unit Root in Time Series Regression. *Biometrika* 75: 335–46. [CrossRef]
- Robe, Michel A., and Jonathan Wallen. 2016. Fundamentals, Derivatives Market Information and Oil Price Volatility. *Journal of Futures Markets* 36: 317–44. [CrossRef]
- Silvennoinen, Annastiina, and Susan Thorp. 2013. Financialization, Crisis and Commodity Correlation Dynamics. *Journal of International Financial Markets Institutions & Money* 24: 42–65. [CrossRef]
- Tang, Ke, and Wei Xiong. 2012. Index Investment and the Financialization of Commodities. *Financial Analysts Journal* 68: 54–74. [CrossRef]
- Zhang, Xinxin, Elie Bouri, Yahua Xu, and Gongqiu Zhang. 2022. The asymmetric relationship between returns and option-implied higher moments: Evidence from the crude oil market. *Energy Economics* 109: 105950. [CrossRef]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

Article

On the Risk Spillover from Bitcoin to Altcoins: The Fear of Missing Out and Pump-and-Dump Scheme Effects

Mehmet Balcilar ^{1,2,*} and Huseyin Ozdemir ³

¹ Department of Economics, Eastern Mediterranean University, Northern Cyprus, Via Mersin 10, Famagusta 99628, Turkey

² Department of Economics, OSTIM Technical University, Ankara 06374, Turkey

³ School of Business, Atilim University, Ankara 06830, Turkey

* Correspondence: mehmet@mbalcilar.net

Abstract: This article examines the asymmetric volatility spillover effects between Bitcoin and alternative coin markets at the disaggregate level. We apply a frequency connectedness approach to the daily data of 11 major cryptocurrencies for the period from 1 September 2017 to 2 March 2022. We try to uncover the existence of the “fear of missing out” psychological effect and “pump-and-dump schemes” in the crypto markets. To do that, we estimate the volatility spillovers from Bitcoin to altcoin and the cryptos’ own risk spillovers during bull and bear markets. The spillover results from Bitcoin to altcoin provide mixed results regarding the presence of this theory for major cryptocurrencies. However, the empirical findings carried out by the cryptos’ own spillover effects fully confirm the existence of a fear-of-missing-out effect and pump-and-dump schemes in all cryptocurrencies except for USDT.

Keywords: asymmetric volatility spillover; bitcoin; altcoin; cryptocurrency; frequency connectedness

JEL Classification: C32; E42; E49; G14; G41



Citation: Balcilar, Mehmet, and Huseyin Ozdemir. 2023. On the Risk Spillover from Bitcoin to Altcoins: The Fear of Missing out and Pump-and-Dump Scheme Effects. *Journal of Risk and Financial Management* 16: 41. <https://doi.org/10.3390/jrfm16010041>

Academic Editor: Kentaro Iwatsubo

Received: 10 December 2022

Revised: 3 January 2023

Accepted: 6 January 2023

Published: 9 January 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Cryptocurrencies have drawn considerable interest from investors, policymakers, and regulators since Bitcoin was created by Satoshi Nakamoto (2008). With their widespread popularity among the public, they have become a popular topic in recent times among academicians, investors, portfolio managers, and regulators. Corbet et al. (2019) divided academic publications on cryptocurrencies into five main categories: bubble dynamics (Corbet et al. 2018b; Vranken 2017), regulation (Böhme et al. 2015; De Filippi 2014; Fletcher et al. 2021), cybercrime (Bernabe et al. 2019; Pinzón and Rocha 2016; Wang et al. 2020), diversification (Chemkha et al. 2021; Urquhart and Zhang 2019), and efficiency (Khuntia and Pattanayak 2018; Nan and Kaizoji 2019; Vidal-Tomás and Ibañez 2018). However, we can extend the list of ranges from return-volume relationships and tail riskiness (Balcilar et al. 2017; Fousekis and Tzaferi 2021) to speculation (Blau 2018; Smaniotto and Neto 2020), as well as return-volatility transmissions between cryptocurrencies and other conventional financial markets (Bouri et al. 2018; Charfeddine et al. 2020; Corbet et al. 2018a).

Unlike other financial assets that are not traded on holidays, the crypto market is open 24 h a day including weekends. Thus, it is important to examine volatility spillover within crypto markets at multiple frequencies (Mensi et al. 2021). Rather than looking through the whole directional risk spillovers among crypto assets, as in other studies (i.e., Brandvold et al. 2015; Ciaian et al. 2018; Fasanya et al. 2021; Katsiampa et al. 2019; Koutmos 2018; Mensi et al. 2021; Sensoy et al. 2021; Yi et al. 2018), we concentrate on the accumulated risk spillovers from Bitcoin to other major alternative coins from the first day to longer periods. Our main motivation for doing so is the fact that price movements

in the crypto market are largely determined by Bitcoin (Corbet et al. 2018a; Kumar et al. 2022; Yi et al. 2018). In addition, we examine the cryptos' own risk spillovers at different frequencies. Indeed, our study is mainly related to two strands of studies in the crypto markets: studies investigating the relationship among crypto markets (Aslanidis et al. 2021; Fousekis and Tzaferi 2021) and studies examining the behavioral characteristics of cryptocurrency investors (Baur and Dimpfl 2018; Wang et al. 2021). This study differs from others in terms of the following aspects: The first aspect is the fact that the main risk coming from Bitcoin price movements to altcoins is neglected by these studies. When this fact is ignored, crypto investors cannot take full advantage of the study results. The second aspect is that we decompose the time domain into various frequencies, but we report the accumulated volatility spillover from Bitcoin to altcoins from the first trading day to the long-term investment horizon. In the third aspect, this study indeed provides different perspectives to test which kinds of investors (informed and uninformed) are dominant in the crypto market for the various altcoins under consideration. Lastly, we use a greater number of cryptos and try to select older major cryptocurrencies¹ with higher market capitalization.

Volatility and volatility spillovers have become hot topics of finance research since the development of the conditional heteroscedasticity models of Engle (1982) and Bollerslev (1987). Subsequently, various generalized autoregressive conditional heteroscedasticity (GARCH) models² have been formulized in the literature after recognizing that volatility propagates asymmetrically (Ang and Chen 2002; Baruník et al. 2017). Later, Cappiello et al. (2006) introduced the asymmetric dynamic conditional correlation DCC (ADCC) specification to account for both multivariate and asymmetries in the conditional variances and the conditional correlations. Based on the realized semivariances proposed by Barndorff-Nielsen et al. (2010), Baruník et al. (2016) propose a way to capture volatility spillovers that are due to bad and good volatility. As an alternative volatility spillover measurement approach, Diebold and Yilmaz (2009, 2012) developed a volatility spillover index (the DY index) based on forecast-error variance decompositions from vector autoregressions (VAR). This technique, however, assumes that the spillover effects among markets are the same across different investment horizons. Still, this assumption fails to model market reality. Baruník and Křehlík (2018) extend the time-domain DY index to the frequency domain to overcome this deficiency. Rather than focusing on frequency responses, this approach is interested in assessing shares of uncertainty in one variable due to shocks with varying persistence levels. In addition, the DY index is better than other ways of measuring volatility, such as multivariate asymmetric GARCH models, because it can measure the direction of the spillover effect in short-, medium-, and long-term financial cycles.

Against this background, the main goal of the present paper is to shed light on the existence of informed traders (or insiders) and uninformed noise traders in the cryptocurrency market. "Fear of missing out" (FOMO) and pump-and-dump schemes have attracted the attention of researchers in cryptocurrency markets (Baur and Dimpfl 2018; Delfabbro et al. 2021; Park and Chai 2020; Wang et al. 2021; Xu and Livshits 2019). FOMO is the fear a trader or investor experiences when they miss out on a potentially profitable investment or trading opportunity in the context of financial markets and trading. The FOMO feeling is most apparent when the value of an asset climbs dramatically in a short period. On the other hand, crypto pump-and-dump schemes occur when conspirators use misleading information to inflate the value of a currency, then sell it for a profit. In this respect, it is wise to assume that the greater the number of uninformed noise traders and fraudsters in a certain altcoin market, the greater the risk of spillovers from bitcoin (or altcoins' prices) to related altcoins. The fact that bitcoin price crashes are followed by other altcoins in the cryptocurrency market confirms this argument.

Our study is strongly related to various studies, such as those by Demir et al. (2021) and Brik et al. (2022), in the cryptocurrency finance literature. Demir et al. (2021) investigated the asymmetric effect of Bitcoin on ETH, XRP, and LTC using the nonlinear autoregressive distributed lag (NARDL) model for the period of July 2015 to March 2019.

Their results indicated that the price of Bitcoin impacts altcoin prices asymmetrically in the short run for all altcoins and a decline in Bitcoin price has a higher effect on altcoins than a rise in Bitcoin price. On the other hand, Brik et al. (2022) examined the return and volatility transmissions between Bitcoin and ten stable and nonstable major cryptocurrencies from 8 October 2018 to 17 August 2020 utilizing the VARMA-BEKK-GARCH model. They provided evidence that volatility transmission is bidirectional in the short and long runs for Bitcoin/Ethereum and Bitcoin/Bitcoin Cash but unidirectional in the short run for Bitcoin/Tether and Bitcoin/TrueUSD. Except for Bitcoin and TrueUSD, there is no long-term bidirectional volatility transmission. The main motivation for this study is to use asymmetric spillover measures to bring a different perspective to the issues that are frequently discussed in the finance literature. It is important for informed investors who want to invest in the crypto market to know in which altcoin market uninformed investors and fraudsters trade the most. Consequently, our work may be of interest to investors who want to recognize the risk coming from uninformed investors and adjust their investment strategy according to BTC price fluctuations.

We examine bitcoin and a set of 10 major altcoins with the largest market capitalization and find interesting results about the spillover from Bitcoin to altcoins during different market conditions. Our data span from 1 September 2017 to 2 March 2022. Our empirical results show that the short- and long-term risk spillovers sourced by Bitcoin are larger for BNB, ETH, LTC, and USDT during bullish market conditions. However, the short- and long-term risk spillovers from Bitcoin to TRON and XRP are greater during bearish market conditions. For ADA and DOGE, the risk spillover emanating from Bitcoin during a bearish market exceeds the risk spillover during a bullish market after the tenth day. Moreover, regarding the risk spread from Bitcoin to BCH and LINK, we do not observe any obvious difference between the bull and bear markets. In addition to this analysis, we examine the cryptocurrency’s own risk spillovers at various frequencies and volatility spillovers from BTC to altcoins during the pre- and post-COVID-19 periods. After COVID-19, the volatility spillover index from BTC to altcoins differs. For instance, the volatility spillover from BTC to BCH, DOGE, LINK, and TRX increased after the COVID-19 outbreak. The empirical findings clearly support FOMO and pump-and-dump schemes for all cryptocurrencies under consideration. Overall, we conclude that the FOMO of noise traders and the deployment of pump-and-dump schemes are inherent features of cryptocurrencies.

The rest of the paper is organized as follows. Section 2 entails the methodology used in this study. Section 3 gives a brief description of the data and some descriptive statistics about the data. Section 4 presents the empirical findings and discussion. Section 5 provides a conclusion.

2. Methodology

This study uses the frequency connectedness approach developed by Baruník and Křehlík (2018) to examine the risk (volatility) spillover from Bitcoin to eight major cryptocurrencies. This technique is an extension of the time-domain spillover index developed by Diebold and Yilmaz (2012). We start by defining the VAR(p) model as

$$x_t = \sum_{m=1}^p \Phi_m x_{t-m} + \varepsilon_t,$$

where $x_t = (x_{1,t}, \dots, x_{N,t})'$ is an N -dimensional covariance stationary stochastic process, Φ_m is an $N \times N$ coefficient matrix, $\varepsilon_t \sim N(0, \Sigma_\varepsilon)$ is an N -dimensional white noise or innovation process, and p is the lag length. Utilizing lag-polynomial approximation (i.e., $\Phi(L) = [I_N - \Phi_1 L - \dots - \Phi_p L^p]$), the VAR model can be written concisely as $\Phi(L)x_t = \varepsilon_t$. The generalized forecast-error variance decomposition (FEVD) proposed by Koop et al. (1996) and Pesaran and Shin (1998), hereafter KPSS, can be computed using the moving average (MA) representation,

$$x_t = \Psi(L)\varepsilon_t,$$

where $\Psi(L)$ stands for infinite lag polynomials and is calculated recursively from $\Psi(L) = \Phi(L)^{-1}$. The H -step-ahead error variances in forecasting x_j are originated from two sources. One of them is due to the corresponding variable's (x_j) own variance, and others are due to the other variable's (x_k) cross variance. Hence, the H -step-ahead generalized FEVD can be calculated as

$$\theta_{j,k}^H = \frac{\sigma_{kk}^{-1} \sum_{h=0}^H \left((\Psi_h \Sigma)_{j,k} \right)^2}{\sum_{h=0}^H (\Psi_h \Sigma \Psi_h')_{j,j}}, \tag{1}$$

where Ψ_h has $N \times N$ -dimensional MA coefficients at lag h and $\sigma_{kk} = (\Sigma)_{k,k}$. $\theta_{j,k}^H$ represents how much of the future forecast error variance of the variable j is due to innovations in variable k at horizon h . Since the rows of the variance decomposition matrix θ^H do not usually sum to one, we need to normalize each entry of the variance decomposition matrix by the row sum as

$$\tilde{\theta}_{j,k}^H = \theta_{j,k}^H / \sum_{j=1}^N \theta_{j,k}^H.$$

Now, $\sum_{k=1}^N \tilde{\theta}_{j,k}^H = 1$ and $\sum_{j,k=1}^N \tilde{\theta}_{j,k}^H = N$ by construction. Using $\tilde{\theta}_{j,k}^H$, we can construct several spillover measures (i.e., total, directional, net, and net pairwise). The total spillover index (S^H) can be constructed using the volatility contributions from the KPPS variance decomposition as

$$S^H = 100 \cdot \frac{\sum_{j,k=1}^N \tilde{\theta}_{j,k}^H}{\sum_{j,k=1}^N \tilde{\theta}_{j,k}^H} = 100 \cdot \frac{\sum_{j,k=1}^N \tilde{\theta}_{j,k}^H}{N}. \tag{2}$$

In addition to the total spillover index, it is possible to measure the spillover transmitted from the overall system to variable j as

$$S_{j \leftarrow \blacksquare}^H = 100 \cdot \frac{\sum_{k=1}^N \tilde{\theta}_{j,k}^H}{N}. \tag{3}$$

Similarly, the spillover transmitted from j to the overall system as

$$S_{j \rightarrow \blacksquare}^H = 100 \cdot \frac{\sum_{k=1}^N \tilde{\theta}_{k,j}^H}{N}. \tag{4}$$

The net spillover index for element j is the difference

$$S^H = S_{\blacksquare \rightarrow j}^H - S_{j \leftarrow \blacksquare}^H. \tag{5}$$

Lastly, the net pairwise spillover between markets j and k is simply the difference between the gross volatility shocks transmitted from market j to market k and those transmitted from k to j as

$$S_{j,k}^H = 100 \cdot \left(\frac{\tilde{\theta}_{k,j}^H - \tilde{\theta}_{j,k}^H}{N} \right). \tag{6}$$

To measure the volatility spillover in the frequency domain, we follow Baruník and Křehlík (2018) and describe the spectral formulation of the variance decomposition. For this purpose, we utilize the Fourier transform of the coefficients Ψ_h (the impulse function used for the time domain) to obtain a frequency response function at a frequency ω

$(\Psi(e^{-i\omega}) = \sum_h e^{-i\omega h} \Psi_h$ with $i = \sqrt{-1}$). The power spectrum $S_X(\omega)$ (Fourier transform of the $MA(\infty)$ filtered series) is

$$S_X(\omega) = \sum_{h=-\infty}^{\infty} E(x_t x'_{t-h}) e^{-i\omega h} = \Psi(e^{-i\omega}) \Sigma \Psi'(e^{+i\omega}). \tag{7}$$

Therefore, Equation (7) illustrates how the variance of the N -dimensional stochastic process is distributed over the frequency components ω . Using the spectral representation for covariance, $E(x_t x'_t) = \int_{-\pi}^{\pi} S_X(\omega) e^{i\omega h} d\omega$, we can define the variance decomposition on the frequency band $d = (\alpha, \beta) : \alpha, \beta \in (-\pi, \pi), \alpha < \beta$ as

$$(\Theta_d)_{j,k} = \frac{1}{2\pi} \int_d \Gamma_j(\omega) \zeta(\omega)_{j,k} d\omega, \tag{8}$$

where $\zeta(\omega)_{j,k}$ is the generalized causation spectrum,

$$\zeta(\omega)_{j,k} = \frac{\sigma_{kk}^{-1} |(\Psi(e^{-i\omega}) \Sigma)_{j,k}|^2}{(\Psi(e^{-i\omega}) \Sigma \Psi'(e^{+i\omega}) \Sigma)_{j,j}},$$

which stands for the portion of the spectrum of the j th variable due to shocks in the k th variable at frequency ω . On the other hand,

$$\Gamma_j(\omega) = \frac{(\Psi(e^{-i\omega}) \Sigma \Psi'(e^{+i\omega}))_{j,j}}{\frac{1}{2\pi} \int_{-\pi}^{\pi} (\Psi(e^{-i\lambda}) \Sigma \Psi'(e^{+i\lambda}))_{j,j} d\lambda'}$$

is a weighting function, and it represents the power of the j th variable at ω , which sums through frequencies to a constant value of 2π . Baruník and Křehlík (2018) termed $(\Theta_d)_{j,k}$ as frequency spillover. After normalizing the values of $(\Theta_d)_{j,k}$ and $\zeta(\omega)_{j,k}$, one can easily calculate alternative spillover indices (Equations (2)–(6)) in the frequency domain.

We define “good” or “bad” volatility³ spillover from Bitcoin to alternative cryptocurrencies when Bitcoin’s daily return⁴ is positive or negative, respectively. This is a good proxy for altcoin investors who observe downside (and upside) risk in the crypto market. Hence, we can formulate good and bad volatility spillovers at a given frequency as follows:

$$S^+(\omega) = r_t > 0, \quad S^-(\omega) = r_t < 0,$$

where $S^+(\omega)$ and $S^-(\omega)$ represent good and bad volatility spillover, and r_t denotes the daily log return of Bitcoin.

3. Data and Descriptive Statistics

We employ daily prices for the eleven major cryptocurrencies (Bitcoin, Cardano, Binance Coin, Bitcoin Cash, Dogecoin, Ethereum, Chainlink, Litecoin, Tron, Tether, and Ripple). The data have daily frequencies and span from 1 September 2017 to 2 March 2022, with the equivalent of 1613 observations. Among them, 861 observations correspond to Bitcoin’s bullish periods and 752 to bearish periods. Furthermore, the World Health Organization (WHO) declared COVID-19 a global epidemic on 11 March 2020. For COVID-19 analysis, we split the whole dataset into two sub-periods (pre- and post-COVID-19) based on this declaration. We take the pre-COVID-19 period before the official declaration and the post-COVID-19 period after that announcement. Moreover, the selection of these cryptocurrencies is motivated by their large market capitalization and long trading periods in comparison to other crypto markets. The cryptocurrencies analyzed in this study account for nearly 75% of total crypto market capitalization.

Figure 1 plots the volatility series of corresponding cryptocurrencies and shows that the volatility series fluctuates over time. Moreover, all volatility series tend to appear in

clusters. In other words, high-volatility sub-periods are followed by low-volatility periods for crypto assets. Moreover, we observe the co-movement of the volatility series, and this can be easily seen in the correlation heat map (see Figure 2). USDT volatility is the least correlated asset with other cryptos, and this is followed by DOGE, among other cryptocurrencies. The volatility of the USDT has also dropped dramatically since the end of 2020. We may argue that the sharp decrease in USDT volatility near the end of 2020 and its stabilization, as a result, are factors that weaken the link between it and other cryptocurrencies. On the other hand, BTC and ETH are the most correlated assets with other cryptos. The descriptive statistics for the crypto volatility series are shown in Table 1. LINK has the highest average volatility, while USDT has the lowest. Moreover, DOGE (USDT) has the greatest (lowest) volatility standard deviation. In addition, Jarque-Bera (JB) reveals that the normality hypothesis is rejected for all volatility series.

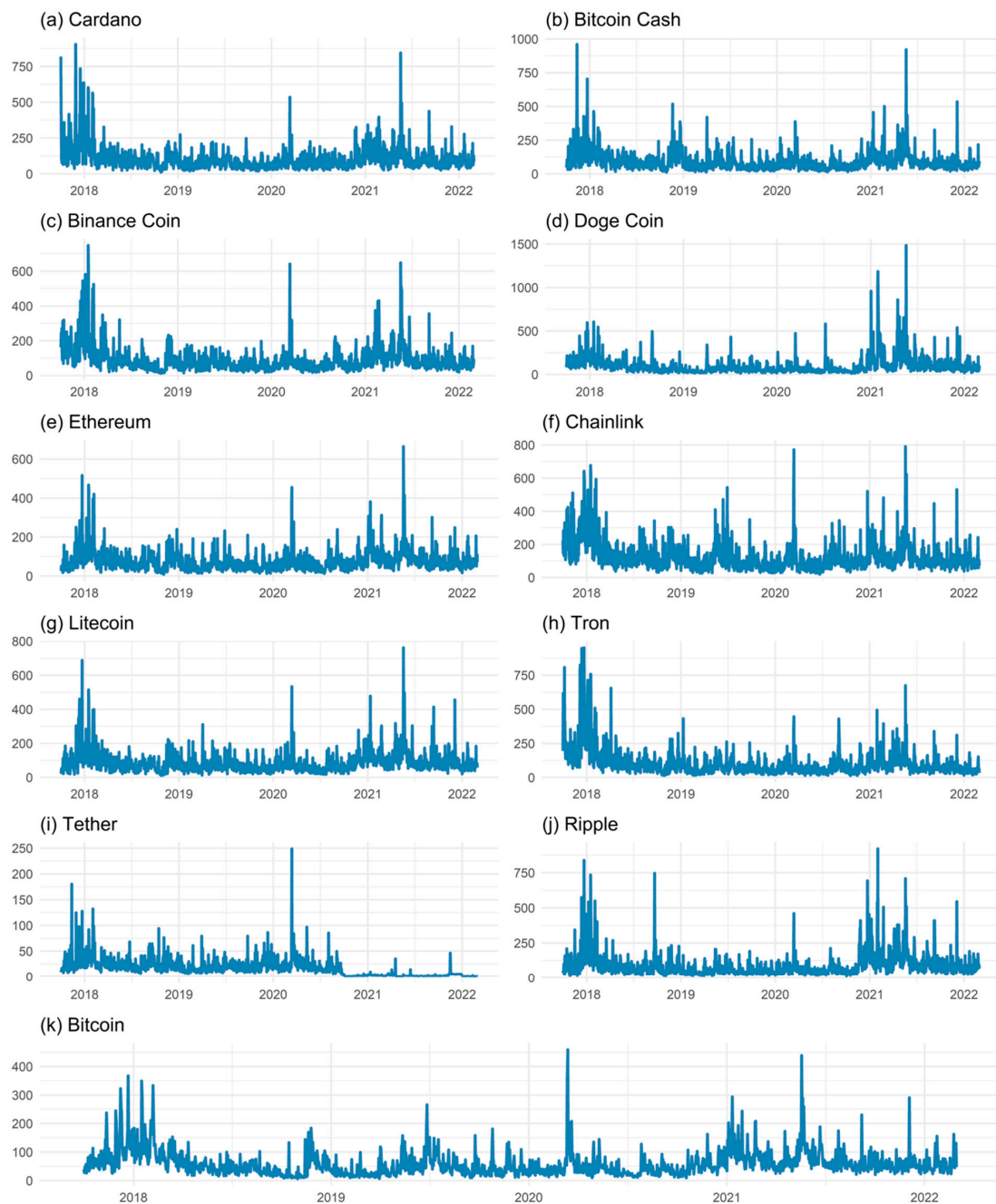


Figure 1. The plot of cryptocurrency volatility series.

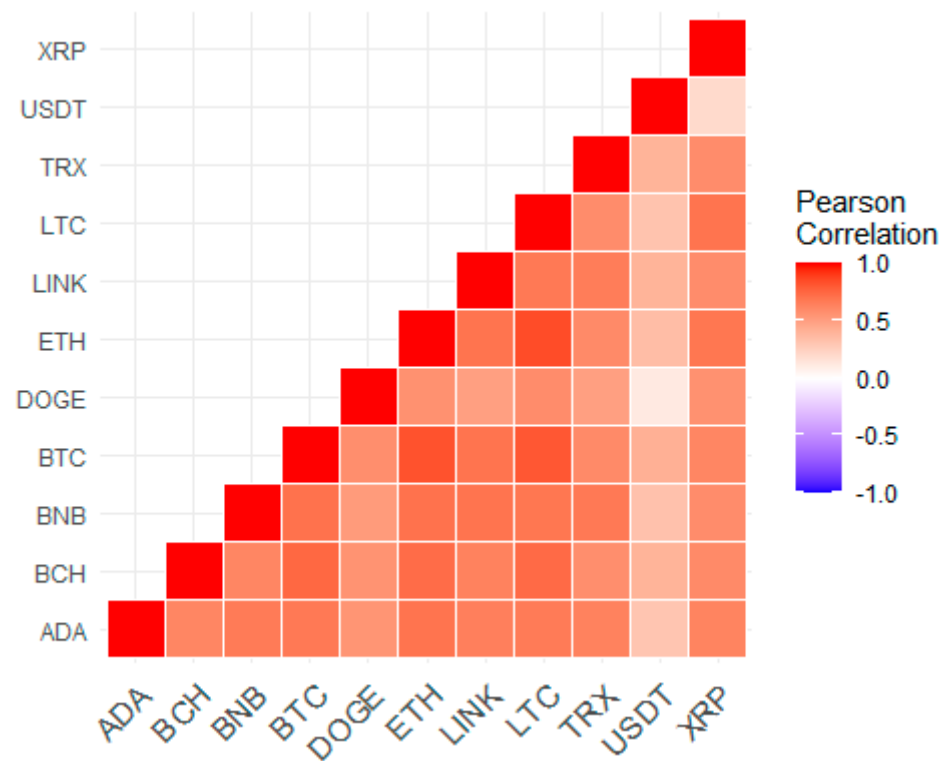


Figure 2. Correlation heat map.

Table 1. Descriptive statistics of volatility series.

	Mean	Max	Min	Std. dev.	Skewness	Kurtosis	Jarque-Bera
ADA	102.77	905.00	11.00	83.68	3.74	25.19	36,860.69 ***
BCH	91.20	961.29	11.96	75.46	3.90	30.42	54,594.35 ***
BNB	90.25	748.61	12.50	74.03	3.35	19.90	22,219.43 ***
DOGE	100.28	1484.94	12.54	107.43	4.62	38.64	91,120.25 ***
ETH	75.16	664.39	7.52	55.11	3.38	23.20	30,504.73 ***
LINK	124.77	791.07	17.72	86.49	2.62	13.61	9415.126 ***
LTC	85.19	762.61	10.11	63.62	3.53	24.63	34,783.51 ***
TRX	103.08	952.12	13.95	99.28	3.66	22.07	28,045.39 ***
USDT	17.05	248.74	0.35	17.97	3.41	29.10	48,899.68 ***
XRP	89.22	920.59	9.96	87.52	3.64	22.65	29,525.00 ***
BTC	58.71	459.05	5.85	46.02	2.85	16.54	14,491.33 ***

Note: This table provides the descriptive statistics of the eleven cryptocurrencies over the period of 1 September 2017–2 March 2022. *** These numbers represent the rejection of the null hypothesis of normality at the 1% significance level.

4. Empirical Results and Discussion

We estimate two VAR models for eleven cryptocurrency volatility series. One is estimated when the Bitcoin return is positive (bullish market), whereas the other is estimated when the Bitcoin return is negative (bearish market). The lag length of both VAR models is determined using the Schwarz Information Criterion (SIC), and H has been set equal to 100. Frequency spillovers are assessed at the bands (0, 1], (1, 2], . . . , (10, 20], (20, 100], and (100, ∞] days. In accordance with Mensi et al. (2021), the short-term ranges between 1 and 5 days, the medium term, 5–20 days, and the long term, more than 20 days. Hence, we calculate the short-, medium-, and long-term risk spillovers from Bitcoin to other altcoins by summing up the volatility spillover in each band (Figure 3).

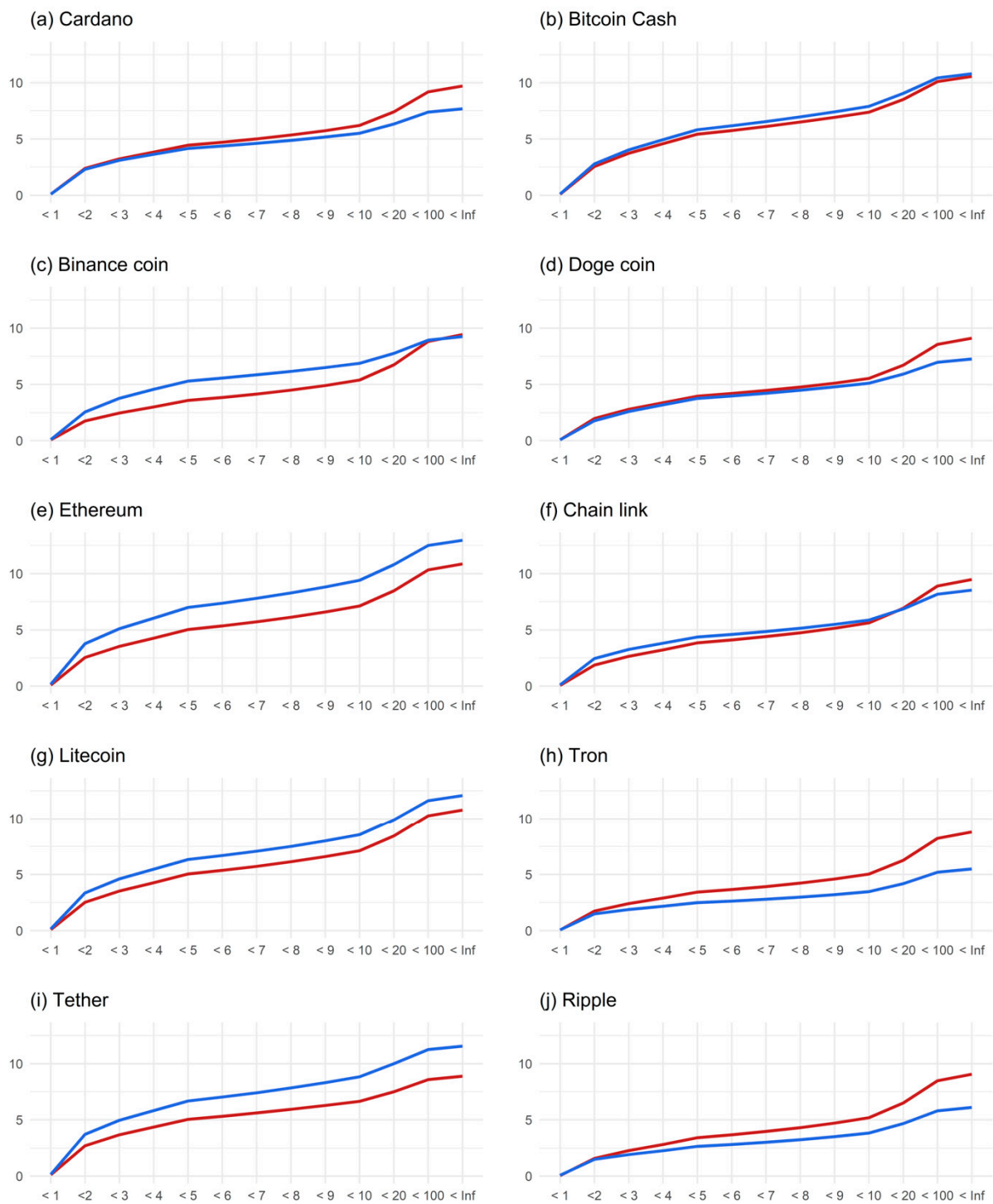


Figure 3. Short- and long-term volatility spillover from Bitcoin to altcoins during bear and bull market conditions. Note: The horizontal axis denotes the frequency in days, while the vertical axis denotes cumulative spillover index. The red and blue line shows the volatility spillover from Bitcoin to altcoins when bitcoin returns are negative (bear market condition) and positive (bull market condition), respectively.

Figure 3 reveals mixed results regarding asymmetry. The short-, medium-, and long-term risk spillover from Bitcoin to BNB, ETH, LTC, and USDT is greater during a bull market than during a bear market. That means BNB, ETH, LTC, and USDT traders are more likely to encounter uninformed investors⁵ compared to other cryptocurrencies analyzed.

Accordingly, these crypto traders should worry more about uninformed investors' herding when making trading decisions because it is difficult to predict when the dump process starts and the overpriced coins will suddenly crash. This means a significant risk of loss of income for these investors. However, for TRON and XRP, the risk spillover is stronger in a bear market. In other words, informed investors are more dominant in these altcoin markets throughout all investment periods. Furthermore, our empirical findings show that the risk spillover from Bitcoin to ADA and DOGE during a bear market exceeds that during a bull market in the medium term. Lastly, the risk spillovers from Bitcoin to BCH and LINK during bull and bear market conditions are very close to each other. Hence, our empirical findings, except for BCH and LINK, are in line with Mensi et al. (2021), who found that cryptocurrencies are sensitive to frequencies.

In addition to the analysis carried out above, we report the asymmetric volatility spillovers of cryptocurrencies (Figure A1, Appendix A) explained by their own shocks. This can be justified by the following reasonable approach: Crypto investors generally follow the price of their invested coin rather than the bitcoin price. Similar to the previous analysis, we consider the period when the bitcoin price increases as a bullish market, whereas negative bitcoin returns represent a bearish market. Figure A1 shows that the risk spillovers of cryptos, except for USDT, on themselves are greater during a bull market than during a bear market. This can be explained by the well-known features of USDT. USDT is a fiat-collateralized stable coin where the actual US dollar currency backs each USDT in circulation. These results strongly confirm Baur and Dimpfl (2018), who found that positive shocks increase volatility more than negative shocks. In other words, the findings are compatible with uninformed investors' FOMO and the presence of pump-and-dump schemes.

The empirical results in Figure 3 also provide evidence about the price formation of altcoins in the short, medium, and long runs. In fact, nearly 70% of the risk spillover from Bitcoin to altcoins takes place in the first 10 days, which can be seen as short-term. This finding concords with the empirical results from earlier studies, such as Ciaian et al. (2018) and Kumar et al. (2022). Among them, Ciaian et al. (2018) found that the prices of altcoins are driven by the development of Bitcoin in the short run but not in the long run. Likewise, Kumar et al. (2022) found that cryptocurrencies are more sensitive to crisis periods over short time horizons than those over longer ones.

We also investigate the spillover effects from BTC to altcoins before and after COVID-19 (Figure A2, Appendix A). This analysis brings a different perspective to the study because incredible price volatility has been observed in cryptocurrency prices since COVID-19. Moreover, during the COVID-19 pandemic, decentralized finance (DeFi) and decentralized apps (dApps) enjoyed a substantial gain in market share and popularity. This outbreak crisis also increased interest in cryptocurrency markets among ordinary and institutional investors, who have access to financial markets from their homes via fintech trading platforms, such as Robinhood (Katsiampa et al. 2022). These facts encourage us to expand this study to examine the asymmetry, including the COVID-19 health crisis. Our empirical findings show that the volatility spillover from BTC to BCH, DOGE, LINK, and TRX increased after the COVID-19 outbreak. Interestingly, ADA and XRP are two cryptocurrencies whose volatility spillover from BTC decreased after the COVID-19 outbreak.

5. Conclusions

In recent years, there has been an increase in empirical research on the fear of missing out (FOMO) on rewarding experiences. There are several studies in the literature that address this issue for stock markets. However, this topic has not yet been sufficiently investigated for cryptocurrencies. This study contributes to filling this gap by analyzing the asymmetric volatility of cryptocurrencies under different market conditions. This study examines the risk spillover from Bitcoin to 10 major cryptocurrencies using daily data from 1 September 2017 to 2 March 2022 and the frequency connectedness analysis of Baruník and Křehlík (2018).

This study attempts to uncover the behaviors of crypto traders by analyzing the risk spillover from Bitcoin to altcoins. Our findings show that BNB, ETH, LTC, and USDT volatility increase more in response to positive shocks than in response to negative shocks in the short, medium, and long terms. This can be explained by uniformed investors' herding, fear of missing out on rising Bitcoin prices, and pump-and-dump schemes. In the short run, this is not the case for ADA and DOGE. For TRON and XRP, we find that the risk spillover is stronger in a bear market. This can be explained by the contrarian behavior of informed investors in these altcoin markets. In addition to this analysis, we investigate the cryptos' own risk spillovers at different frequencies. These empirical findings strongly support the FOMO and pump-and-dump schemes for all cryptocurrencies except for USDT. Moreover, we employ the volatility spillover effects from BTC to altcoins, considering the COVID-19 crisis. The empirical findings show that the risk from BTC to ADA and XRP decreases after the post-covid period in comparison to the pre-pandemic period.

The results of this paper suggest that the impact of Bitcoin price movements on altcoin prices is mixed, meaning that the relationship between the two is not always predictable. This can make it difficult for investors to use Bitcoin price movements as a reliable indicator of altcoin performance. However, the presence of FOMO and pump-and-dump activity in the altcoin market suggests that individual altcoin prices may be more susceptible to manipulation and volatility spillover. As such, investors may want to be cautious when considering investments in altcoins and may want to thoroughly research the market and individual cryptocurrencies before making any investment decisions. It may also be beneficial for investors to diversify their portfolios by including a mix of Bitcoin and altcoins rather than putting all their eggs in one basket. Moreover, investors should also keep track of the latest developments in the cryptocurrency market, including any news or regulatory changes that could impact the market. This can help them stay informed and make informed investment decisions. Investors should also be wary of the fear of missing out and not make impulsive investment decisions based on hype or media coverage. Investors should also use reputable exchanges because they are less likely to engage in manipulative behavior or facilitate pump-and-dump schemes.

We point out that the results of this study are based on broad market analysis based on global data. The evidence we provide for the existence of fear of missing out and pump-and-dump schemes is limited to the altcoin market. Therefore, the psychological factors play a lesser role for cryptos with large market capitalization, such as Bitcoin. Moreover, our analysis results are valid for the broad market and may not apply to all individual exchanges.

Given that the price movements of altcoins are highly dependent on BTC and their own price history, future studies may extend our approach to analyze volatility spillover from other leading cryptocurrencies, such as ETH, ADA, and SOL, to various altcoins related to these platforms. Polygon, formerly known as Matic Network, for example, is a Layer-2 scaling solution for Ethereum that aims to enhance the network's transaction processing speed while lowering transaction costs, often known as "gas prices." In that respect, one may investigate the risk spillover from ETH to MATIC when the ETH price is going up and down. Another extension is to investigate the higher frequency (hour or minute) volatility connectedness of cryptocurrency to uncover FOMO and pump-and-dump strategies for the benefit of high-frequency traders.

Author Contributions: Conceptualization, M.B. and H.O.; methodology, H.O.; software, H.O.; validation, M.B. and H.O.; formal analysis, H.O.; investigation, H.O.; resources, H.O.; data curation, H.O.; writing—original draft preparation, M.B. and H.O.; writing—review and editing, M.B. and H.O.; visualization, M.B. and H.O. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

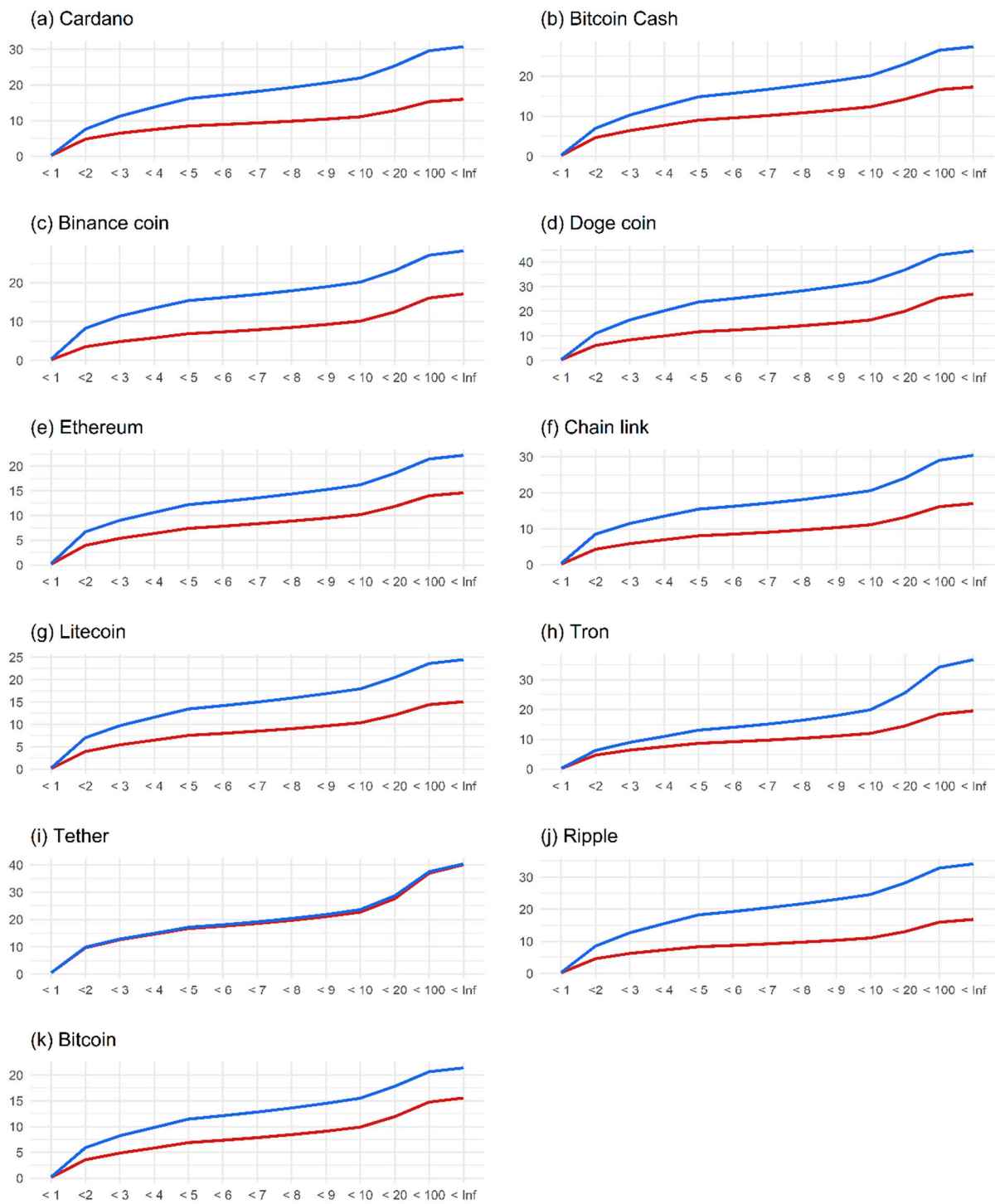


Figure A1. Short- and long-term volatility spillover from altcoins to themselves. Note: See note to Figure 3.

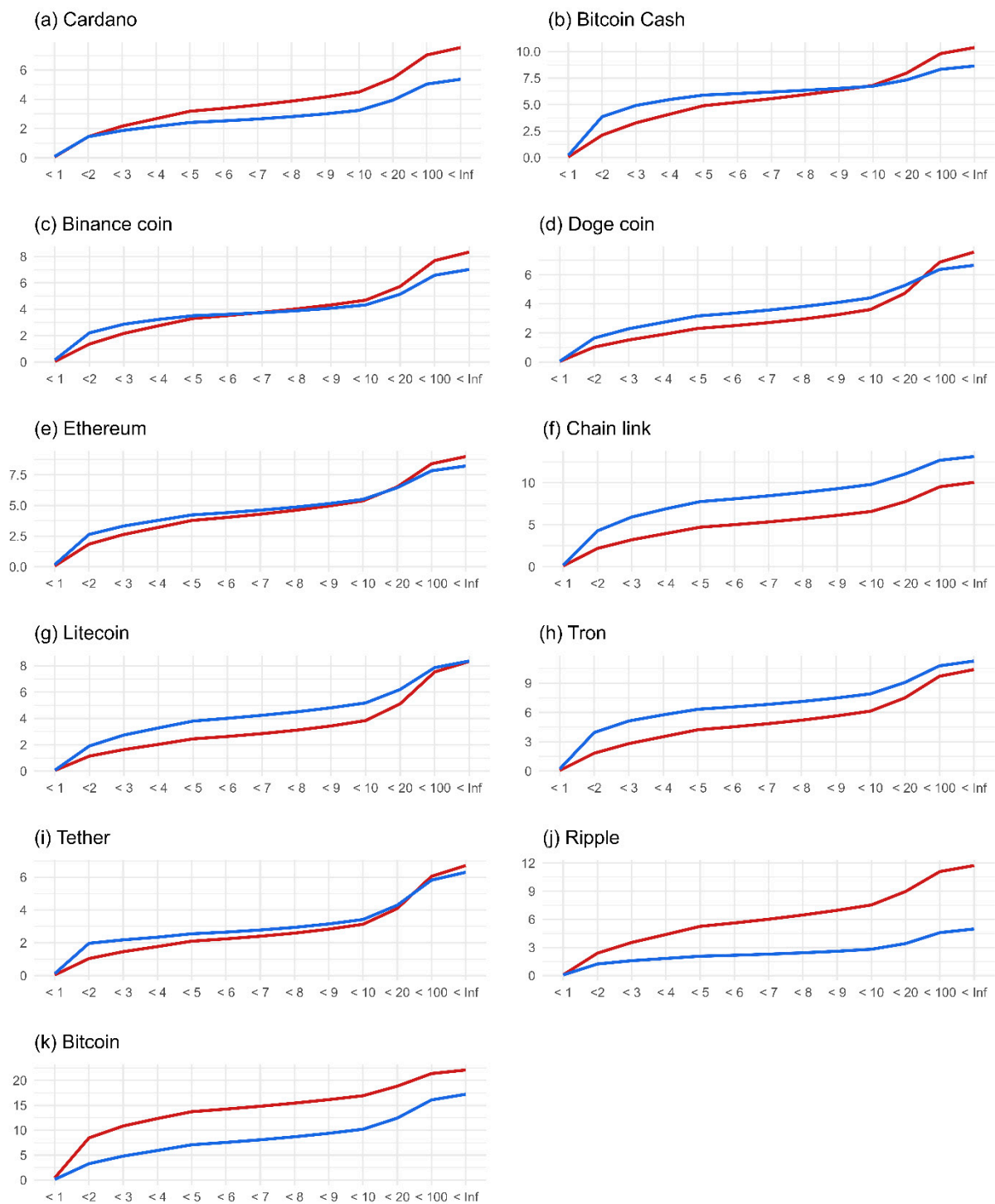


Figure A2. Short- and long-term volatility spillover from Bitcoin to altcoins during pre- and post-COVID-19 period. Note: The horizontal axis denotes the frequency in days, while the vertical axis denotes cumulative spillover index. The red and blue line shows the volatility spillover from Bitcoin to altcoins during pre-COVID-19 and post-COVID-19 period, respectively.

Notes

- ¹ Bitcoin (BTC), Cardano (ADA), Binance Coin (BNB), Bitcoin Cash (BCH), Dogecoin (Doge), Ethereum (ETH), Chainlink (LINK), Litecoin (LTC), Iron (TRX), Tether (USDT), and Ripple (XRP).
- ² See, e.g., Nelson (1991) for E-GARCH, Glosten et al. (1993) for leverage effect GARCH, and (Zakoian 1994) for TGARCH.

- 3 Following Garman and Klass (1980), we calculate the daily volatility series using the daily opening, closing, high, and low prices. Firstly, the daily volatility is calculated as $\tilde{\sigma}_{it}^2 = \frac{1}{2} [\ln(P_{it}^{\max}) - \ln(P_{it}^{\min})]^2 - [2 \ln(2) - 1] [\ln(P_{it}^{\text{close}}) - \ln(P_{it}^{\text{open}})]^2$, where P_{it}^{\max} , P_{it}^{\min} , P_{it}^{close} , and P_{it}^{open} show the minimum, the maximum, the close, and the opening price of the market i on day t , respectively. Second, we annualize the volatility series utilizing the formula $\hat{\sigma}_{it} = 100\sqrt{365 \times \tilde{\sigma}_{it}^2}$.
- 4 We calculate the daily returns of Bitcoin by taking the logarithm of the close price divided by the open price as $r_t = 100 \cdot \ln\left(\frac{P_t^{\text{close}}}{P_t^{\text{open}}}\right)$. The daily observations span from 1 September 2017, to 2 March 2022.
- 5 The behavior of informed investors is consistent with the basic suggestions of economic theory. They put more emphasis on investment knowledge and economic-related criteria than uninformed investors, who are more influenced by behavioral elements such as personality and sentiment (Jalilvand et al. 2018). Based on this fact, it is fair to think that uninformed investors are more open to market rumors with no economic justification.

References

- Ang, Andrew, and Joseph Chen. 2002. Asymmetric correlations of equity portfolios. *Journal of Financial Economics* 63: 443–94. [CrossRef]
- Aslanidis, Nektarios, Aurelio F. Bariviera, and Alejandro Perez-Laborda. 2021. Are cryptocurrencies becoming more interconnected? *Economics Letters* 199: 109725. [CrossRef]
- Balcilar, Mehmet, Elie Bouri, Rangan Gupta, and David Roubaud. 2017. Can volume predict Bitcoin returns and volatility? A quantiles-based approach. *Economic Modelling* 64: 74–81. [CrossRef]
- Barndorff-Nielsen, Ole E., Silja Kinnebrock, and Neil Shephard. 2010. Measuring Downside Risk—Realized Semivariance. In *Volatility and Time Series Econometrics: Essays in Honor of Robert Engle*. Oxford: Oxford Academic. [CrossRef]
- Baruník, Jozef, and Tomáš Křehlík. 2018. Measuring the frequency dynamics of financial connectedness and systemic risk. *Journal of Financial Econometrics* 16: 271–96. [CrossRef]
- Baruník, Jozef, Evžen Kočenda, and Lukáš Vácha. 2016. Asymmetric connectedness on the U.S. stock market: Bad and good volatility spillovers. *Journal of Financial Markets* 27: 55–78. [CrossRef]
- Baruník, Jozef, Evžen Kočenda, and Lukáš Vácha. 2017. Asymmetric volatility connectedness on the forex market. *Journal of International Money and Finance* 77: 39–56. [CrossRef]
- Baur, Dirk G., and Thomas Dimpfl. 2018. Asymmetric volatility in cryptocurrencies. *Economics Letters* 173: 148–51. [CrossRef]
- Bernabe, Jorge Bernal, Jose Luis Canovas, Jose L. Hernandez-Ramos, Rafael Torres Moreno, and Antonio Skarmeta. 2019. Privacy-Preserving Solutions for Blockchain: Review and Challenges. *IEEE Access* 7: 164908–164940. [CrossRef]
- Blau, Benjamin M. 2018. Price dynamics and speculative trading in Bitcoin. *Research in International Business and Finance* 43: 493–99. [CrossRef]
- Böhme, Rainer, Nicolas Christin, Benjamin Edelman, and Tyler Moore. 2015. Bitcoin: Economics, technology, and governance. *Journal of Economic Perspectives* 29: 213–38. [CrossRef]
- Bollerslev, Tim. 1987. A Conditionally Heteroskedastic Time Series Model for Speculative Prices and Rates of Return. *The Review of Economics and Statistics* 69: 542–47. [CrossRef]
- Bouri, Elie, Mahamitra Das, Rangan Gupta, and David Roubaud. 2018. Spillovers between Bitcoin and other assets during bear and bull markets. *Applied Economics* 50: 5935–49. [CrossRef]
- Brandvold, Morten, Peter Molnár, Kristian Vagstad, and Ole Christian Andreas Valstad. 2015. Price discovery on Bitcoin exchanges. *Journal of International Financial Markets, Institutions and Money* 36: 18–35. [CrossRef]
- Brik, Hatem, el Ouakdi Jihene, and Ftiti Zied. 2022. Roles of stable versus nonstable cryptocurrencies in Bitcoin market dynamics. *Research in International Business and Finance* 62: 101720. [CrossRef]
- Cappiello, Lorenzo, Robert F. Engle, and Kevin Sheppard. 2006. Asymmetric dynamics in the correlations of global equity and bond returns. *Journal of Financial Econometrics* 4: 537–72. [CrossRef]
- Charfeddine, Lanouar, Noureddine Benlagha, and Youcef Maouchi. 2020. Investigating the dynamic relationship between cryptocurrencies and conventional assets: Implications for financial investors. *Economic Modelling* 85: 198–217. [CrossRef]
- Chemkha, Rahma, Ahmed BenSaïda, Ahmed Ghorbel, and Tahar Tayachi. 2021. Hedge and safe haven properties during COVID-19: Evidence from Bitcoin and gold. *Quarterly Review of Economics and Finance* 82: 71–85. [CrossRef]
- Ciaian, Pavel, Rajcaniova Miroslava, and Kancs d'Artis. 2018. Virtual relationships: Short- and long-run evidence from BitCoin and altcoin markets. *Journal of International Financial Markets, Institutions and Money* 52: 173–95. [CrossRef]
- Corbet, Shaen, Andrew Meegan, Charles Larkin, Brian Lucey, and Larisa Yarovaya. 2018a. Exploring the dynamic relationships between cryptocurrencies and other financial assets. *Economics Letters* 165: 28–34. [CrossRef]
- Corbet, Shaen, Brian Lucey, and Larisa Yarovaya. 2018b. Datestamping the Bitcoin and Ethereum bubbles. *Finance Research Letters* 26: 81–88. [CrossRef]
- Corbet, Shaen, Brian Lucey, Andrew Urquhart, and Larisa Yarovaya. 2019. Cryptocurrencies as a financial asset: A systematic analysis. *International Review of Financial Analysis* 62: 182–99. [CrossRef]
- De Filippi, Primavera. 2014. Bitcoin: A regulatory nightmare to a libertarian dream. *Internet Policy Review* 3. [CrossRef]
- Delfabbro, Paul, Daniel L. King, and Jennifer Williams. 2021. The psychology of cryptocurrency trading: Risk and protective factors. *Journal of Behavioral Addictions* 10: 201–7. [CrossRef]

- Demir, Ender, Serdar Simonyan, Conrado-Diego García-Gómez, and Chi Keung Marco Lau. 2021. The asymmetric effect of bitcoin on altcoins: Evidence from the nonlinear autoregressive distributed lag (NARDL) model. *Finance Research Letters* 40: 101754. [CrossRef]
- Diebold, Francis X., and Kamil Yilmaz. 2009. Measuring financial asset return and volatility spillovers, with application to global equity markets. *Economic Journal* 119: 158–71. [CrossRef]
- Diebold, Francis X., and Kamil Yilmaz. 2012. Better to give than to receive: Predictive directional measurement of volatility spillovers. *International Journal of Forecasting* 28: 57–66. [CrossRef]
- Engle, Robert F. 1982. Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation. *Econometrica* 50: 987–1007. [CrossRef]
- Fasanya, Ismail Olaleke, Oluwatomisin Oyewole, and Temitope Odudu. 2021. Returns and volatility spillovers among cryptocurrency portfolios. *International Journal of Managerial Finance* 17: 327–41. [CrossRef]
- Fletcher, Emily, Charles Larkin, and Shaen Corbet. 2021. Countering money laundering and terrorist financing: A case for bitcoin regulation. *Research in International Business and Finance* 56: 101387. [CrossRef]
- Fousekis, Panos, and Dimitra Tzaferi. 2021. Returns and volume: Frequency connectedness in cryptocurrency markets. *Economic Modelling* 95: 13–20. [CrossRef]
- Garman, Mark B., and Michael J. Klass. 1980. On the Estimation of Security Price Volatilities from Historical Data. *The Journal of Business* 53: 67–78. [CrossRef]
- Glosten, Lawrence R., Ravi Jagannathan, and David E. Runkle. 1993. On the Relation between the Expected Value and the Volatility of the Nominal Excess Return on Stocks. *The Journal of Finance* 48: 1779–801. [CrossRef]
- Jalilvand, Abolhassan, Mojtaba Rostami Noroozabad, and Jeannette Switzer. 2018. Informed and uninformed investors in Iran: Evidence from the Tehran Stock Exchange. *Journal of Economics and Business* 95: 47–58. [CrossRef]
- Katsiampa, Paraskevi, Larisa Yarovaya, and Damian Zieba. 2022. High-Frequency connectedness between bitcoin and other top-traded crypto assets during the COVID-19 crisis. *Journal of International Financial Markets, Institutions and Money* 79: 101578. [CrossRef]
- Katsiampa, Paraskevi, Shaen Corbet, and Brian Lucey. 2019. High frequency volatility co-movements in cryptocurrency markets. *Journal of International Financial Markets, Institutions and Money* 62: 35–52. [CrossRef]
- Khuntia, Sashikanta, and J. K. Pattanayak. 2018. Adaptive market hypothesis and evolving predictability of bitcoin. *Economics Letters* 167: 26–28. [CrossRef]
- Koop, Gary, M. Hashem Pesaran, and Simon M. Potter. 1996. Impulse response analysis in nonlinear multivariate models. *Journal of Econometrics* 74: 119–47. [CrossRef]
- Koutmos, Dimitrios. 2018. Return and volatility spillovers among cryptocurrencies. *Economics Letters* 173: 122–27. [CrossRef]
- Kumar, Ashish, Najaf Iqbal, Subrata Kumar Mitra, Ladislav Kristoufek, and Elie Bouri. 2022. Connectedness among major cryptocurrencies in standard times and during the COVID-19 outbreak. *Journal of International Financial Markets, Institutions and Money* 77: 101523. [CrossRef]
- Mensi, Walid, Khamis Hamed Al-Yahyaee, Idries Mohammad Wanas Al-Jarrah, Xuan Vinh Vo, and Sang Hoon Kang. 2021. Does volatility connectedness across major cryptocurrencies behave the same at different frequencies? A portfolio risk analysis. *International Review of Economics and Finance* 76: 96–113. [CrossRef]
- Nakamoto, Satoshi. 2008. Bitcoin: A Peer-to-Peer Electronic Cash System. Available online: <https://bitcoin.org/bitcoin.pdf> (accessed on 10 December 2022).
- Nan, Zheng, and Taisei Kaizoji. 2019. Market efficiency of the bitcoin exchange rate: Weak and semi-strong form tests with the spot, futures and forward foreign exchange rates. *International Review of Financial Analysis* 64: 273–81. [CrossRef]
- Nelson, Daniel B. 1991. Conditional Heteroskedasticity in Asset Returns: A New Approach. *Econometrica* 59: 347–70. [CrossRef]
- Park, Minjung, and Sangmi Chai. 2020. The effect of information asymmetry on investment behavior in cryptocurrency market. Paper presented at the Annual Hawaii International Conference on System Sciences, Maui, HI, USA, January 7–10.
- Pesaran, H. Hashem, and Yongcheol Shin. 1998. Generalized impulse response analysis in linear multivariate models. *Economics Letters* 58: 17–29. [CrossRef]
- Pinzón, Carlos, and Camilo Rocha. 2016. Double-spend Attack Models with Time Advantage for Bitcoin. *Electronic Notes in Theoretical Computer Science* 329: 79–103. [CrossRef]
- Sensoy, Ahmet, Thiago Christiano Silva, Shaen Corbet, and Benjamin Miranda Tabak. 2021. High-frequency return and volatility spillovers among cryptocurrencies. *Applied Economics* 53: 4310–28. [CrossRef]
- Smaniotta, Emanuelle Nava, and Giacomo Balbinotto Neto. 2020. Speculative trading in Bitcoin: A Brazilian market evidence. *Quarterly Review of Economics and Finance* 85: 47–54. [CrossRef]
- Urquhart, Andrew, and Hanxiong Zhang. 2019. Is Bitcoin a hedge or safe haven for currencies? An intraday analysis. *International Review of Financial Analysis* 63: 49–57. [CrossRef]
- Vidal-Tomás, David, and Ana Ibañez. 2018. Semi-strong efficiency of Bitcoin. *Finance Research Letters* 27: 259–65. [CrossRef]
- Vranken, Harald. 2017. Sustainability of bitcoin and blockchains. *Current Opinion in Environmental Sustainability* 28: 1–9. [CrossRef]
- Wang, Jying-Nan, Hung-Chun Liu, Shuang Zhang, and Yuan-Teng Hsu. 2021. How does the informed trading impact Bitcoin returns and volatility? *Applied Economics* 53: 3223–33. [CrossRef]
- Wang, Qin, Bo Qin, Jiankun Hu, and Fu Xiao. 2020. Preserving transaction privacy in bitcoin. *Future Generation Computer Systems* 107: 793–804. [CrossRef]

- Xu, Jiahua, and Benjamin Livshits. 2019. The anatomy of a cryptocurrency pump-and-dump scheme. Paper presented at the 28th USENIX Security Symposium, Santa Clara, CA, USA, August 14–16.
- Yi, Shuyue, Zishuang Xu, and Gang-Jin Wang. 2018. Volatility connectedness in the cryptocurrency market: Is Bitcoin a dominant cryptocurrency? *International Review of Financial Analysis* 60: 98–114. [CrossRef]
- Zakoian, Jean-Michel. 1994. Threshold heteroskedastic models. *Journal of Economic Dynamics and Control* 18: 931–55. [CrossRef]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

Article

Analysis of Bitcoin Price Prediction Using Machine Learning

Junwei Chen 

Graduate School of Economics, Kobe University, Kobe 657-8501, Japan; junweichen1110@gmail.com

Abstract: The research purpose of this paper is to obtain an algorithm model with high prediction accuracy for the price of Bitcoin on the next day through random forest regression and LSTM, and to explain which variables have influence on the price of Bitcoin. There is much prior literature on Bitcoin price prediction research, and the research methods mainly revolve around the ARMA model of time series and the LSTM algorithm of deep learning. Although it cannot be proved by the Diebold–Mariano test that the prediction accuracy of random forest regression is significantly better than that of LSTM, the prediction errors RMSE and MAPE of random forest regression are better than those of LSTM. The changes in the variables that determine the price of Bitcoin in each period are also obtained through random forest regression. From 2015 to 2018, three US stock market indexes, NASDAQ, DJI, and S&P500 and oil price, and ETH price have impact on Bitcoin prices. Since 2018, the important variables have become ETH price and Japanese stock market index JP225. The relationship between accuracy and the number of periods of explanatory variables brought into the model shows that for predicting the price of Bitcoin for the next day, the model with only one lag of the explanatory variables has the best prediction accuracy.

Keywords: Bitcoin; machine learning; random forest regression; LSTM

1. Introduction

Bitcoin is a decentralized digital currency that uses cryptography for security and is not controlled by any government or financial institution. It was created in 2008 by an individual or group of individuals using the pseudonym Satoshi Nakamoto (2008) with a paper titled “Bitcoin: A Peer-to-Peer (P2P) Electronic Cash System”. Transactions with bitcoin are recorded on a public ledger called the blockchain, which allows anyone to view the history of a specific Bitcoin. The decentralized nature of Bitcoin allows it to operate independently of central banks and can be transferred instantly across the globe. It has gained popularity as a means of exchange and a store of value (Baur and Dimpfl 2021). In the past 10 years, after experiencing several ups and downs, it broke through USD 68,000 per coin in November 2021, and the total current price once exceeded USD 1.2 trillion.

However, as a commodity, Bitcoin has the problem of high volatility. During the seven years from April 2015 to April 2022, the standard deviation of Bitcoin’s daily return rate was 3.85%, which was 2.68 times the standard deviation of gold’s return rate during the same period and 3.36 times that of the S&P500. Due to the large price fluctuations, the function of Bitcoin as a store of value as a commodity and as a transaction payment function as a currency has been questioned.

While enjoying the advantages of Bitcoin’s security and decentralization, how to grasp the trend of Bitcoin to minimize the risk of Bitcoin floating has become a difficult problem. Many researchers try to grasp the trend of Bitcoin through the correlation between the price of Bitcoin and the price of other commodities. But whether it is gold (Baur and Hoang 2021; Kim et al. 2020b; Blake 2019), which is often used for comparison, stock market index (Erdas and Caglar 2018), or crude oil price (Selmi et al. 2018), past studies have shown that the correlation between Bitcoin and them is weak.

In past studies, another type of research direction to grasp the price trend of Bitcoin is to predict the price of Bitcoin in the future through AI algorithms and powerful computing



Citation: Chen, Junwei. 2023. Analysis of Bitcoin Price Prediction Using Machine Learning. *Journal of Risk and Financial Management* 16: 51. <https://doi.org/10.3390/jrfm16010051>

Academic Editor: Kentaro Iwatsubo

Received: 28 December 2022

Revised: 10 January 2023

Accepted: 11 January 2023

Published: 13 January 2023



Copyright: © 2023 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

power of computers. With the improvement of hardware performance in the 21st century, machine learning technology which has become a hot field of research. Primarily, machine learning has been used across a variety of areas such as that of stock markets (Huang and Liu 2020; Philip 2020); crude oil markets (Fan et al. 2016); gold markets (Chen et al. 2020b); and futures markets (Kim et al. 2020a).

Prediction of Bitcoin by AI is mainly divided into two categories. The first category is the classification research of predicting the rise or fall of Bitcoin in the future. The error standard is DA and F1. The other category is regression research on predicting Bitcoin prices, while the corresponding errors are RMSE and MAPE. Due to the sharp fluctuations in the price of Bitcoin, only grasping the rise or fall of the price of Bitcoin in the future cannot help investors avoid risks. In contrast, getting the specific bitcoin price as a reference price is more useful.

1.1. Motivation and Novelties

Based on the necessity of avoiding the price risk of Bitcoin as the background, this research chooses the random forest regression algorithm of machine learning and the LSTM model of neural network algorithm to predict the price of Bitcoin. I mainly focus on the performance of random forest regression in Bitcoin price prediction when using the prediction results of LSTM as a comparison. Random forest regression is a regression form of random forest. Different from the black box technology of neural networks, random forest regression as machine learning can deliver the importance of each explanatory variable in predicting Bitcoin through the results of its weak-learners.

The prediction effect of random forest in predicting stock price direction has been proven effective (Basak et al. 2019; Khan et al. 2020). However, unlike random forest classifier, whose research goal is to classify ups or downs, there are not many papers that use random forest regression to study the cryptocurrency market in the existing literature. In the literature using random forest regression, the explanatory variables used by Parvez (2022) focus on the highly correlated OHLC (Open, High, Low, Close) and transaction volume of Bitcoin itself as explanatory variables. On this basis, I think it is of great research value to add explanatory variables in other fields. A total of 47 explanatory variables were collected for this study in the following 8 categories: (a) Bitcoin price variables, (b) the specific technical features of Bitcoin, (c) other cryptocurrencies, (d) commodities, (e) market index, (f) foreign exchange, (g) public attention, and (h) dummy variables of the week to verify the accuracy of random forest regression for Bitcoin price prediction.

As a comparison of whether the prediction accuracy of random forest regression is good, this paper chooses the LSTM algorithm of RNN as comparative research. The experimental results of many studies show that the prediction accuracy of LSTM and GRU is better when compared with other models, including the traditional time series model ARMA.

In addition to pursuing a high-precision forecasting model, this study also conducts (1) an in-depth analysis from the explanatory variables that determine the importance of Bitcoin prices and (2) the relationship between the prediction accuracy and the lag of the explanatory variables.

1.2. Contributions

The RMSE of the random forest regression model is smaller than LSTM algorithm. Although through the DM and Clark–West test, the hypothesis that LSTM is better than random forest regression cannot be rejected at a significant level of $\alpha = 95\%$. However, the error results of multiple experiments show the higher prediction accuracy of random forest regression.

The experimental results of random forest regression also indicate the changes in the factors that determine the price of Bitcoin around 2018. The OHLC prices of Bitcoin itself are proven to be most important during the full sample period. In Period 1 from April 2015 to October 2018, the U.S. stock markets NASDAQ, DJI, and S&P500, which have high

importance, show a sharp decrease in importance in the Period 2 sample from October 2018 to April 2022. The importance of ETH and DOGE, which are both digital currency markets, increased during Period 2.

As an LSTM model that focuses on the study of time series data, the control experiments by substituting explanatory variables with different lags show that the prediction accuracy obtained only with the latest period of data is the highest. Random forest regression also delivered the same conclusion.

1.3. Organization

Rest of the paper is organized as follows. Section 2 discusses the existing methodologies and models to predict the cryptocurrency prices. Section 3 discusses the setting of model parameters and error setting. Section 4 discusses the selection analysis and pre-processing of explanatory variables. Section 5 discusses the performance evaluation of the proposed model. Section 6 discusses the limitations of the research and directions for future attempts. Finally, Section 7 concludes the paper.

2. Related Works

Aggarwal et al. (2019) studied whether gold price can predict Bitcoin price through three deep learning algorithms of CNN, LSTM, and GRU. The conclusion is that the predicted price of the model which only uses gold price deviates from the true Bitcoin price, and the prediction accuracy of the LSTM model is the best of three. Liu et al. (2021) expanded the range of explanatory variables, based on the cryptocurrency market and macro market index (stock market index, crude oil price, exchange rate, etc.) and search index, a total of 40 explanatory variables for Bitcoin price prediction. SDAE algorithm shows better prediction performance than BPNN, PCA-SVR, and SVR.

Regarding the prediction research of Bitcoin price, the methods are divided into time series and machine learning. Multiple studies have concluded that the prediction accuracy of ARIMA is not as good as that of machine learning (McNally et al. 2018; Shin et al. 2021; Chen et al. 2020a; Akyildirim et al. 2021).

LSTM, as a controlled study of random forest regression in this study, has been studied as a target model many times in the past literature (Shin et al. 2021; Jagannath et al. 2021; Rizwan et al. 2019). Phaladisailoed and Numnonda (2018) used four deep learning algorithms (Theil–Sen regression, Huber regression, LSTM, and GRU) to predict the price of Bitcoin. The 52.78% accuracy of the LSTM algorithm is the highest. Based on the same explanatory variables, Tandon et al. (2019) found that adding 10-fold cross-validation to the LSTM training process can increase the accuracy of LSTM by 14.7%. However, the selection of explanatory variables in Phaladisailoed's and Tandon's studies is limited to OHLC, volume from top exchange and market cap. In the research done by Aggarwal et al. (2019), in addition to the price of Bitcoin itself, gold price was added to explanatory variables. The experimental results show that the RMSE of the LSTM algorithm is 47.91, which is better than CNN and GRU. McNally et al. (2018) added the variables difficulty and hash rate related to Bitcoin attributes in his research, the 52.78% prediction accuracy of LSTM is also better than the accuracy of RNN and ARIMA. Chen et al. (2020a) used LSTM, SVR, ANFIS, and ARIMA, four algorithms to predict the Bitcoin price. While Chen added eight kinds of Bitcoin attribute variables, public attention variables (Google Trends and Twitter data) and economic category variables. In the four subsample periods, LSTM all showed better prediction accuracy than the other three. Livieris et al. (2020) introduced a novel framework by preprocessing, which performed a series of transformations based on first differences or returns, to make data "suitable" for fitting a deep learning model based on the stationarity property.

In addition to predicting the price of Bitcoin, there are many studies using LSTM to predict other digital currencies (Sebastião and Godinho 2021; Saadah and Whafa 2020; Derbentsev et al. 2020). Politis et al. (2021) used LSTM to predict the price of Ether with an accuracy of 84.2%. Livieris et al. (2021) used hybrid CNN-LSTM to conduct prediction

experiments on Bitcoin (BTC), Ethereum (ETH), and Ripple (XRP) with the highest market value at the time and obtained BTC The prediction accuracy of 55.03% is higher than ETH's 51.51% and XRP's 49.61%.

In McNally et al.'s (2018), García-Medina and Duc Huynh's (2021), and Chen et al.'s (2020a) studies, it is mentioned that adding Dropout layers between each layer of LSTM can reduce the effect of overlearning. But there are differences in the choice of dropout coefficients (0.1, 0.3, 0.5) among the three works of literature above.

Regarding the selection of explanatory variables, in addition to the macroeconomic variables used in many works of literature, Jagannath et al.'s (2021) research focuses on the core variables of the Bitcoin blockchain, including users, miners, and exchanges. Technical indicators have proven useful for predicting Bitcoin prices (Jaquart et al. 2021; Mudassir et al. 2020). The LSTM based on the self-adaptive technique also gets good prediction performance, but the article lacks a comparative experiment with the model added macroeconomic variables. Regarding the explanatory power of variables on Bitcoin price, García-Medina and Duc Huynh (2021) innovatively studied variables such as social media (E. Musk and D. Trump's remarks) and Tesla stock price. During the ups and downs in the second half of 2020, the conclusion was that the explanatory power of these variables that were of great interest at the time was not found. Carbó and Gorjón (2022), in their appendix, compare the effect of adding the previous period's Bitcoin price to the explanatory variables based on the LSTM algorithm. The RMSE accuracy of the model that added the previous Bitcoin price as an explanatory variable improved significantly from the original 21% to 11%.

The selection of time unit prices is also a point that has been analyzed by many researchers. Most research use days or minutes as the sample unit. In the quarterly research of DSVR, DNDT, and DRCNN conducted by Lamothe-Fernández et al. (2020), each model obtained more than 60% prediction accuracy, but this high accuracy may be related to Bitcoin's general uptrend between 2011 and 2019 in the sample, as well as the long quarterly units. The work of Shin et al. (2021) is based on the LSTM model, with sample units in a minute, hour, and day. The results show that the prediction accuracy of the day model and minute model is similar, and both better than the model with an hour unit.

Bitcoin has a history of 15 years since its birth in 2008, although it is not long compared to other assets. In previous studies, researchers are more willing to subdivide data samples into small samples before conducting prediction research (Shin et al. 2021; Chen et al. 2020a; Carbó and Gorjón 2022). In Jagannath et al.'s (2021) and Awoke et al.'s (2021) experiments, the longest period of a single sample does not exceed 4 years.

3. Methodology

Machine learning is an important branch of artificial intelligence (AI). According to whether there is a target variable, it can be divided into supervised learning, unsupervised learning, and reinforcement learning. The purpose of this study is to predict future Bitcoin prices, so a regression function with supervised learning is used. The unified execution logic of machine learning is that after the algorithm is preset, a learner is generated, and a high-precision learner is obtained by repeated training of the learner through training data and the process of validation. Finally, the test data is substituted into the trained learner for evaluation and application.

Both random forest regression and LSTM model training in this paper are implemented through the open-source library of python's machine learning. The library used by random forest regression is sklearn, and LSTM uses keras for research. The pre-processing and collation of the data are done by pandas.

3.1. Random Forest

Random forest is an ensemble form of multiple regression trees. Its advantage is high explicability, but the predicted results are limited by the training samples. The principle

of the regression tree is to divide the parent group into subgroups using an indicator of a certain variable, and the classification is based on making the average of the sum of squared residuals of each group the smallest, shown in Equation (1) below.

$$\frac{1}{n_1} \sum_{i=1}^{n_1} (y_i - \bar{y}_{(1:n_1)})^2 + \frac{1}{n_2 - n_1} \sum_{j=n_1+1}^{n_2} (y_j - \bar{y}_{(n_1+1:n_2)})^2 \rightarrow \min \quad (1)$$

Regarding parameter settings, the maximum depth of a single sub-regression tree is 10, and the number of sub-regression trees in the random forest is 500 (Figure 1). I tested the maximum depth of the interval [min = 3, max = 20] and the number of sub-regression trees of the interval [min = 200, max = 1000], respectively. My further experiments show that after the maximum depth is greater than 10 or the number of sub-regression trees is greater than 500, the training data and the prediction error no longer changes.

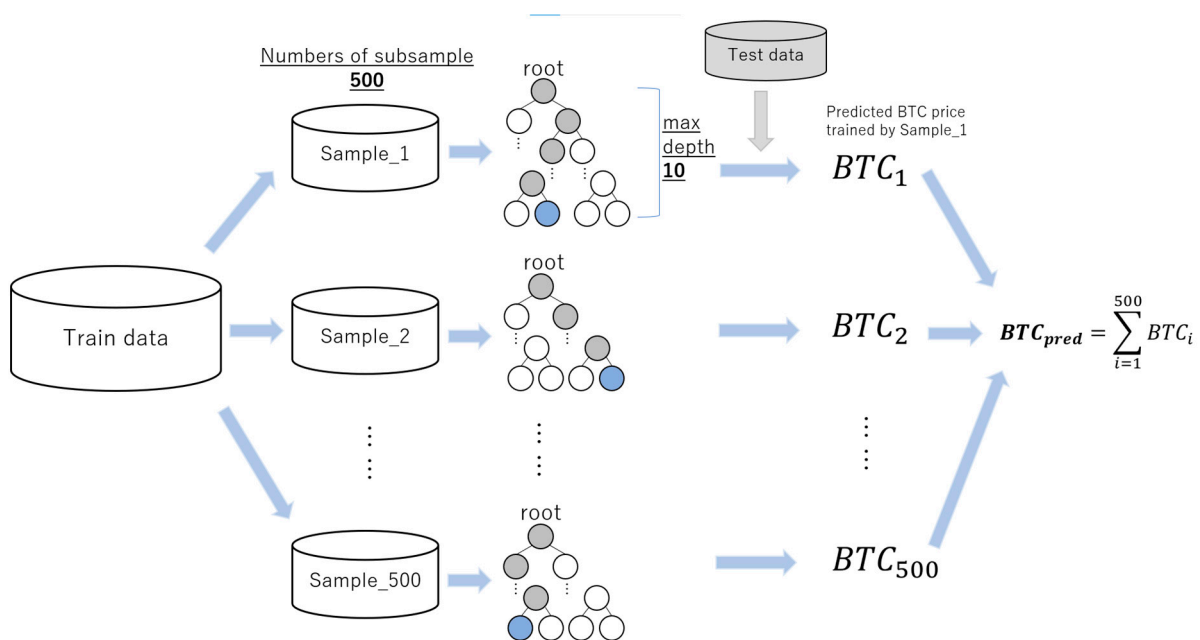


Figure 1. Parameters and framework of random forest regression.

3.2. LSTM

The RNN algorithm is different from the normal DNN algorithm. When data is substituted into the model, it will not only generate an output value, but also modify the parameters of the model. RNN algorithm has the function of retaining the previous input data information in the model. This paper uses the LSTM model that makes up for the short memory defect of RNN. Data changes are made to the RNN model and the memory model through the paths of the three activation functions of Forget Gate, Input Gate, and Output Gate.

Based on the characteristic that the output value of the LSTM model can be re-substituted into another layer of the LSTM model, and the application of the dropout layer mentioned in the literature, the LSTM model structure of this experiment is as follows. Regarding the parameter setting of the dropout layer, I tested [min = 10%, max = 50%] for each dropout layer. It turns out that when the overall value of dropout is small, there is an overlearning phenomenon in which the training data performs well but the prediction error of validation data is large. When the overall value of dropout is set too large, the errors of the training data and the validation data are both large. In addition, the experiment also found that the prediction accuracy of the dropout value with descending order is worse than ascending order. The number of layers of LSTM [min = 2, max = 6] and the parameter setting of each layer of units in [32, 64, 128, 256, 512] are tested. After balancing the accuracy

and the risk of overlearning. The activation function of each layer is set to “ReLU”, which has better performance than “sigmoid” and “tanh”. The specific value and framework of LSTM is shown by Table 1 and Figure 2 below. The last 10% of the training data is set as validation data.

Table 1. Details of the LSTM model.

	Layers	Parameters
Layer_1	LSTM_1	units: 128 Activation: ReLU
Layer_2	Dropout_1	0.2
Layer_3	LSTM_2	units: 128 Activation: ReLU
Layer_4	Dropout_2	0.3
Layer_5	LSTM_3	units: 256 Activation: ReLU
Layer_6	Dropout_3	0.4
Layer_7	LSTM_4	units: 256 Activation: ReLU
Layer_8	Dropout_4	0.5
Layer_9	Dense	units: 1

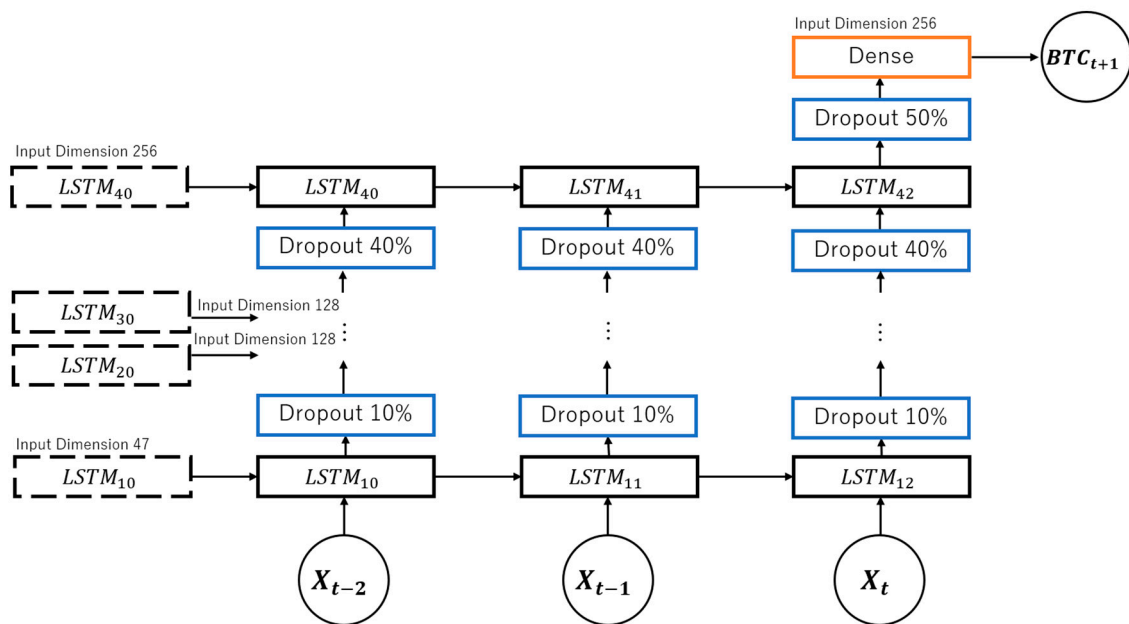


Figure 2. Parameters and framework of LSTM.

In addition to the framework setting of the model, another important hyper parameter of deep learning is epochs. The value of epochs reflects the number of passes to learn the train data. The larger the epochs are, the smaller the prediction error of the training data will be. However, when the epochs are too large, it leads to overlearning problems. Therefore, through the training and validation loss diagrams of Period 1 and Period 2 in Figure 3 below, the epochs of Period 1 are set to 250, and the epochs of Period 1 are set to 75.

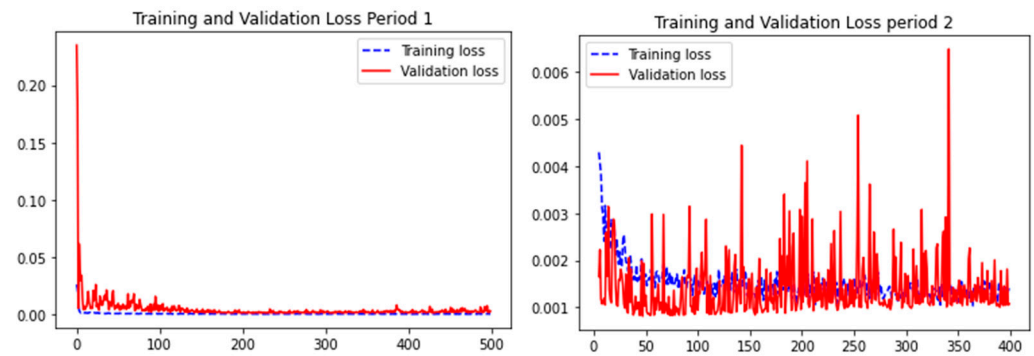


Figure 3. Training and validation loss of LSTM.

3.3. Errors and Evaluation Criteria

As an important criterion to evaluate the prediction accuracy of machine learning, this study quantifies the prediction performance of the model by using three errors, MAPE (mean absolute percentage error, Equation (2)) and RMSE (root mean squared error, Equation (3)), and DA (decision accuracy, Equation (4)). However, due to the rising average Bitcoin price, RMSE can only be compared for model results based on the same sample. There is no meaningful comparison between the experimental results of different data samples.

$$MAPE = \frac{1}{m} \sum_{t=1}^m \left| \frac{y(t) - \hat{y}(t)}{y(t)} \right| \tag{2}$$

$$RMSE = \sqrt{\frac{1}{m} \sum_{t=1}^m (y(t) - \hat{y}(t))^2} \tag{3}$$

$$DA = \frac{1}{m} \sum_{t=1}^m a(t) \times 100\% \tag{4}$$

In addition to comparing the prediction accuracy of various models to obtain the performance of each model in predicting the future price of Bitcoin, this study also expects to compare the prediction errors under different lags of explanatory variables to analyze the memory length characteristics of the Bitcoin market.

In addition to the MAPE, RMSE, and DA errors of each prediction result, this paper also conducts a hypothesis test on the significant difference between the two different algorithms through the Diebold–Mariano test and the Clark–West test. The principle of the DM test can be simply summarized as: given two sets of prediction error sequences $\{e'_t\}_{t=j}^T$ and $\{e_t\}_{t=j}^T$, then define a loss function $d_t = L(e_t) - L(e'_t)$, while $L(e) = e^2$ is MSE and $L(e) = |e|$ is MAE.

$$DM_t = \frac{\bar{d}_t}{se(d_t)} \tag{5}$$

Based on Diebold–Mariano’s loose assumption, DM_t (Equation (5)) is asymptotically distributed in $N(0, 1)$, and finally a one-sided hypothesis test is performed on the statistic DM_t .

The Clark–West test adds the $(e_t - e'_t)^2$ item in the loss function of the Diebold–Mariano test of MSE as $f_t := (e_t)^2 - (e'_t)^2 + (e_t - e'_t)^2$, which is also asymptotically distributed in $N(0, 1)$, and finally performs a one-tailed hypothesis test on the statistic f_t .

4. Data and Preprocessing

The sample data are the daily data from 31 March 2015 to 1 April 2022. The data of the study were collected from yahoo finance, Coinmarketcap.com, investing.com, bitinfocharts.com, and coinmetrics.io.

The target variable in the experiment is the price of Bitcoin in USD. A total of 47 variables are used as explanatory variables to predict the price of Bitcoin in the future, which are divided into eight categories: (a) Bitcoin price variables, (b) the specific technical features of Bitcoin, (c) other cryptocurrencies, (d) commodities, (e) market index, (f) foreign exchange, (g) public attention, and (h) dummy variables of the week.

Each explanatory variable and its corresponding definition are in Appendix A.

4.1. Explanatory Variables Analysis

Table 2 shows the statistical features for each explanatory variable used to predict Bitcoin’s future price. It is worth noting that the standard deviations of the variables related to the cryptocurrency market (five for Bitcoin, five for other cryptocurrencies, and Google search volume for Bitcoin) are all large. Among them, the ratio of the standard deviation to the mean value, except for the LTC of 0.99, all the others exceed 1. It reflects the high volatility of the cryptocurrency market since 2015. Except for the variables mentioned above, which are related to cryptocurrency, the value of standard deviation/mean ratio of the traditional market is not greater than 0.4.

Table 2. Statistical features of explanatory variables.

	Count	Mean	Std	Min	Max
BTC_Open	2559	12,628.14	16,689.78	210.068	67,549.73
BTC_High	2559	12,965.49	17,133.74	223.833	68,789.63
BTC_Low	2559	12,259.05	16,184.48	199.567	66,382.06
BTC_Close	2559	12,644.27	16,697.06	210.495	67,566.83
BTC_Volume	2559	1.6×10^{10}	2.02×10^{10}	10,600,900	3.51×10^{11}
Active addr cnt	2559	715,123	235,979.6	222,628	1,366,494
Xfer cnt	2559	646,493.3	183,825.9	234,806	2,041,653
Mean Tx size (native units)	2559	2.092273	3.50753	0.307039	126.7199
Total fees (USD)	2559	936,734.4	1,971,955	2850.355	21,397,763
Mean hash rate	2559	60,571,448	61,550,129	271,738.1	2.48×10^8
Difficulty	2559	8.37×10^{12}	8.5×10^{12}	4.67×10^{10}	2.86×10^{13}
Mean block size (in bytes)	2559	968,516.6	258,456.1	292,929.3	1,523,656
Sum block weight	2559	4.82×10^8	1.05×10^8	1.91×10^8	7.58×10^8
LTC	2559	71.87075	70.81633	1.32117	386.4508
XRP	2559	0.354487	0.38141	0.00356	2.78
DASH	2559	142.1313	182.4392	2.06	1550.85
DOGE	2559	0.035873	0.087754	8.73×10^{-5}	0.6848
ETH	2430	708.8693	1107.578	0.4348	4812.09
Gold	1854	1489.887	245.8335	1070.8	2117.1
Silver	2182	19.18016	3.750716	11.978	30.135
Copper	1811	3.00615	0.697527	1.994	4.9375
Oil	1848	54.88971	14.53394	-37.63	123.7
Treasury yield 10 years	1763	1.950953	0.657184	0.499	3.234
S&P500	1766	2907.096	779.8341	1829.08	4796.56
DJI	1766	24,828.27	5703.945	15,660.18	36,799.65

Table 2. Cont.

	Count	Mean	Std	Min	Max
CBOE	1765	94.984	21.60072	55.5	137.16
NASDAQ	1765	8336.731	3308.791	4266.84	16,057.44
JP225	1740	21,972.35	3738.272	14,952.02	30,670.1
CSI300	1708	3982.53	668.6175	2853.76	5807.72
DXY	1764	95.63923	2.961022	88.59	103.29
EUR	1826	1.343444	0.088168	1.149439	1.588512
GBP	1826	0.747414	0.046768	0.62952	0.86999
JPY	1826	111.051	5.136474	99.906	125.629
CAD	1826	1.303631	0.04442	1.1954	1.4578
AUD	1826	1.367315	0.07251	1.232	1.741281
SGD	1826	1.367216	0.029435	1.30659	1.4563
CNY	1826	0.733329	0.037271	0.57429	0.811688
RUB	1826	66.58596	8.731132	0.7162	138.9651
Tweets	2559	50,500.83	43,438.57	13,294	363,566
Google	2559	495.8206	519.2102	64	6064.504

In addition, differences between the explanatory variables of the cryptocurrency market and the traditional market were observed in terms of the ratio of the minimum and maximum values. Except for 194 times the Russian ruble in traditional markets, the max/min ratio is not greater than 7 (Regardless of the extremely negative price of -37.63 for crude oil on 20 April 2020). However, in the cryptocurrency market, the ratios are all greater than 300, and the highest is 11,067 times that of ETH. Both the Bitcoin market and the Russian ruble in the traditional market have shown high volatility.

The correlation heat map (Figure 4) shows the correlation between Bitcoin and other explanatory variables except for the week dummy variables. Bitcoin has a positive correlation with other cryptocurrencies, commodity prices, stock market indexes, and public attention variables. The only exception is that the price of Bitcoin is inversely correlated with the 10-year U.S. Treasury yield in the commodities category. The price of Bitcoin and the exchange rate generally show a negative correlation. It seems understandable that the stronger the US dollar, the lower the price of Bitcoin. Interestingly, the Russian ruble exchange rate has a positive correlation with the Bitcoin price, and the correlation coefficient is high.

There is a brief explanation of the relationship between Bitcoin price and weekday variables. The extreme floats are mostly found on Wednesdays. The largest yield variance was seen on Wednesday and the largest daily gains and daily losses over the 7 years both occurred on Wednesday. The variance of yields is smaller on weekends than on weekdays, and yield fluctuations are more stable. The average daily return for Bitcoin is 0.28% with a 95% confidence interval of [0.13%, 0.43%]. The average return is highest on Mondays and smallest on Sundays. Monday’s return is statistically greater than Sunday’s ($\alpha = 95\%$). The daily probability of rising is 54.57% with a 95% confidence interval of [52.64%, 56.50%]. Saturday and Friday have the highest probability of rising. The probability of rising on Saturday is statistically greater than on Sunday ($\alpha = 90\%$).

Regarding the two public attention variables (Figure 5), two conclusions can be drawn from the comparison with the Bitcoin price. First, the spike in Google Trends and daily Tweets came during a time when Bitcoin broke its all-time high price. Secondly, the highest Google Trend occurred at the end of 2017. After that, even with over USD 60,000 in 2021, the search volume did not surpass what it was at the end of 2017.

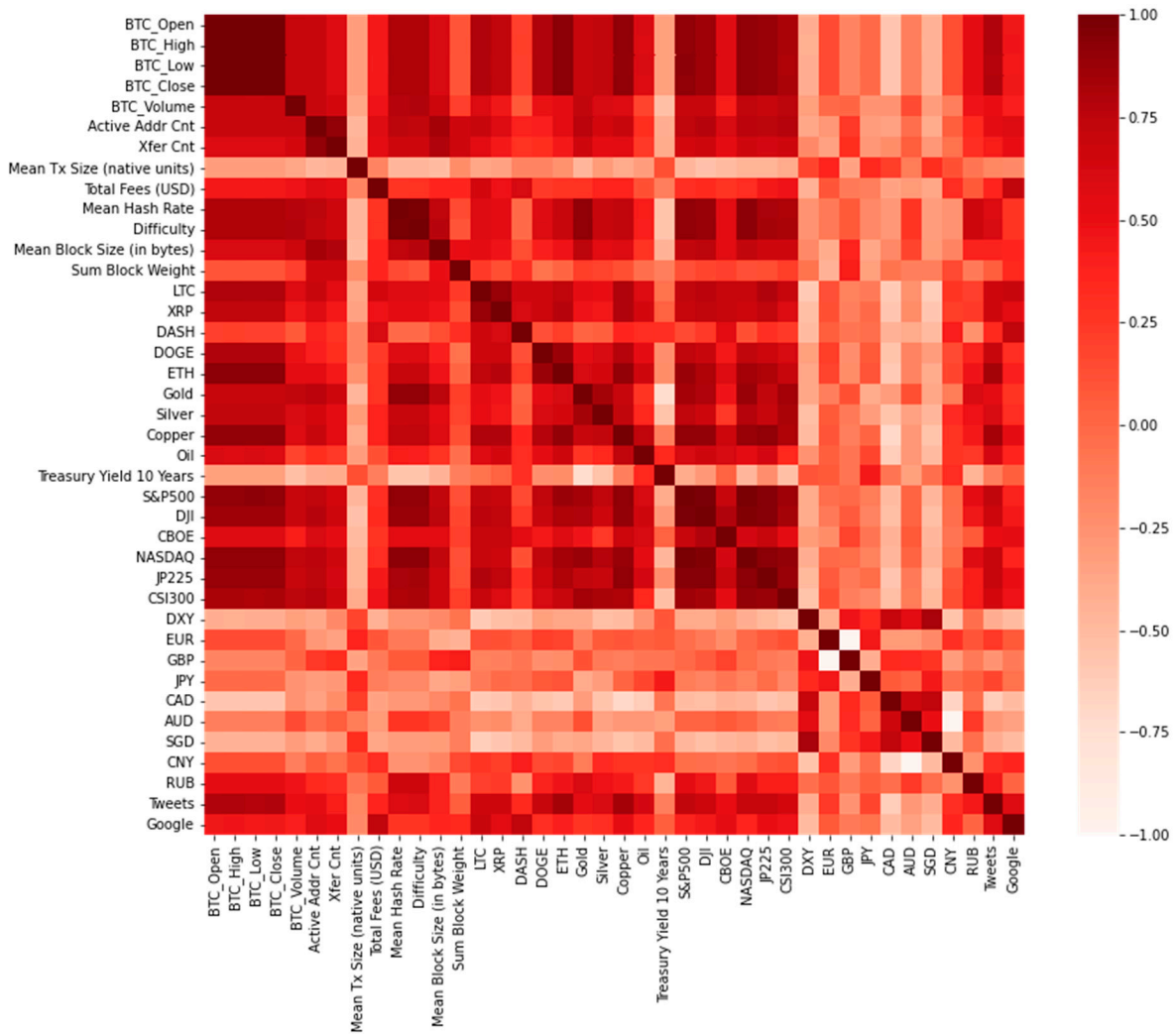


Figure 4. Correlation heatmap of explanatory variables.

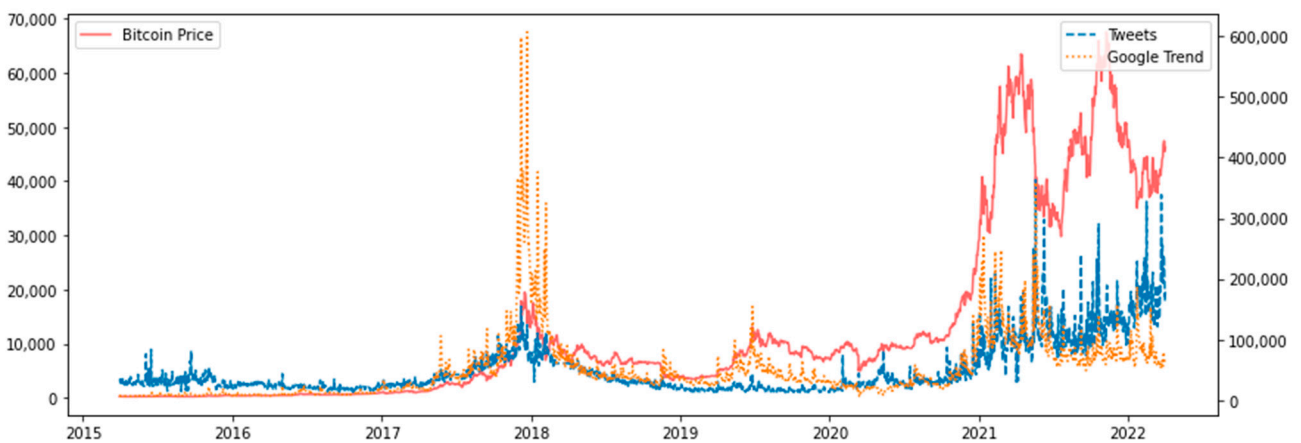


Figure 5. Google Trend, daily Tweets, and Bitcoin price chart.

4.2. Preprocessing

The data research sample collected data from a total of 7 natural years from 31 March 2015 to 1 April 2022. However, due to the particularity of Bitcoin having two price bubbles at the end of 2017 and 2021, and the longest span of a single sample in past studies is no longer

than 4 years. Based on the above two reasons, to improve the price prediction accuracy of the model, the total sample is divided into Period 1 (from 31 March 2015 to 30 September 2018) and Period 2 (1 October 2018 with 1 October 2018). Conduct independent research on two sub-samples, train models for their respective periods and predict respectively. Machine learning is the process of training initial samples through training samples and then substituting them into test samples for evaluation. Usually, training samples occupy 75% to 90% of the samples. The specific division of training and testing samples in this study is shown in Table 3 and Figure 6. The last 10% of the training data is set as validation data.

Table 3. Interval division of training samples and test samples.

	Train Data	Test Data	Percentage of Train Data
Period 1	31 March 2015–31 March 2018	1 April 2018–30 September 2018	85.70%
Period 2	1 October 2018–30 September 2021	1 October 2021–1 April 2022	85.69%

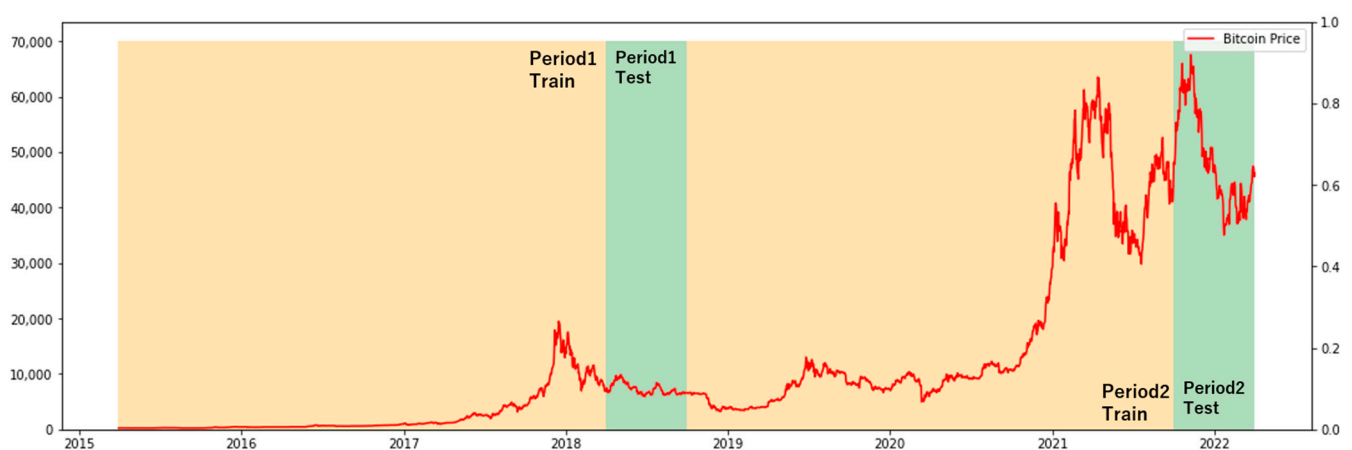


Figure 6. Interval division of training samples and test samples.

Among all the explanatory samples, only ETH has the problem of missing sample data because it came out (7 August 2015) later than April 2015, so the training samples used for ETH in the Period 1 model all start from 7 August 2015, not 31 March 2015.

Bitcoin is available for trading 24 h a day and 365 days a year, while the variables such as stock market indices, exchange rates, and commodity price indices are not traded during weekends and holidays, so there is missing data. There are two ways to deal with samples with these missing data, one is to delete the data with missing data before training, and the second method is to fill in the missing data. Considering that the research object of this study is time-series in nature, direct deletion of the samples affects the analysis of the period relationship. Therefore, filling in the missing data is chosen by replacing the value of the missing data with the value of the previous period. For example, in the case of gold prices, there is no price data for the weekend, and the value of the Friday gold price from the previous day is used to define the price for both days of the weekend.

The min/max preprocessing (Equation (6)) is important for LSTM because the activation function is not sensitive to values above 1. All variables are unified to [0, 1], eliminating the effect of metric units.

$$x_{scaled} = \frac{x - x_{min}}{x_{max} - x_{min}} \tag{6}$$

The flow of the whole experiment is shown in Figure 7.

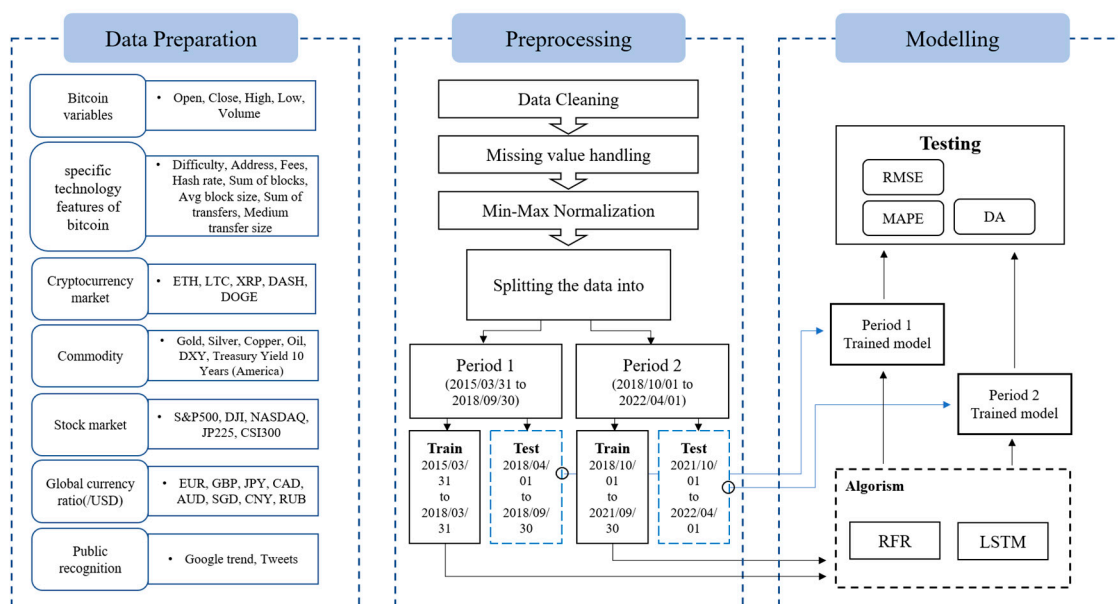


Figure 7. Model employed in this study.

5. Results

5.1. Results of Random Forest Regression

The trained learner is used to predict the test samples of Period 1 and Period 2, and the results shown in Table 4 and Figure 8 below are obtained. The red line is the Bitcoin price, and the green dashed line is the price predicted by the random forest regression learner.

Table 4. Error results for random forest regression.

	Period 1	Period 2
RMSE	321.61	2096.24
MAPE	3.39%	3.29%
DA	51.93%	52.49%

Although the RMSE of Period 1 is much smaller than that of Period 2, since the average price of Bitcoin in Period 1 is also much smaller than that of Period 2, it is meaningless to compare the RMSE results of different periods. The MAPE and DA indicators in the two periods are quite close, and the prediction accuracy of Period 2 is slightly better than that of Period 1. It is worth noting that in the early stage of the test interval of Period 2, the random forest regression algorithm has a bad prediction on the Bitcoin price when the price is greater than 60,000 US dollars because there are very few samples with a Bitcoin price greater than USD 60,000 in the training samples of Period 2. This result accurately reflects the disability of the random forest algorithm to predict results outside the training samples. However, whether it is Period 1 or Period 2, the random forest regression algorithm shows excellent performance in predicting prices below USD 60,000, and the trend of the predicted price is consistent with the real price trend.

In addition to predictive analysis, the random forest algorithm also provides the importance of each explanatory variable when predicting the price of Bitcoin, through the statistics of the number of occurrences of boundary variables in all 500 sub-regression trees. The result is shown in Figure 9.



Figure 8. Predicted price based on random forest regression and actual price comparison.

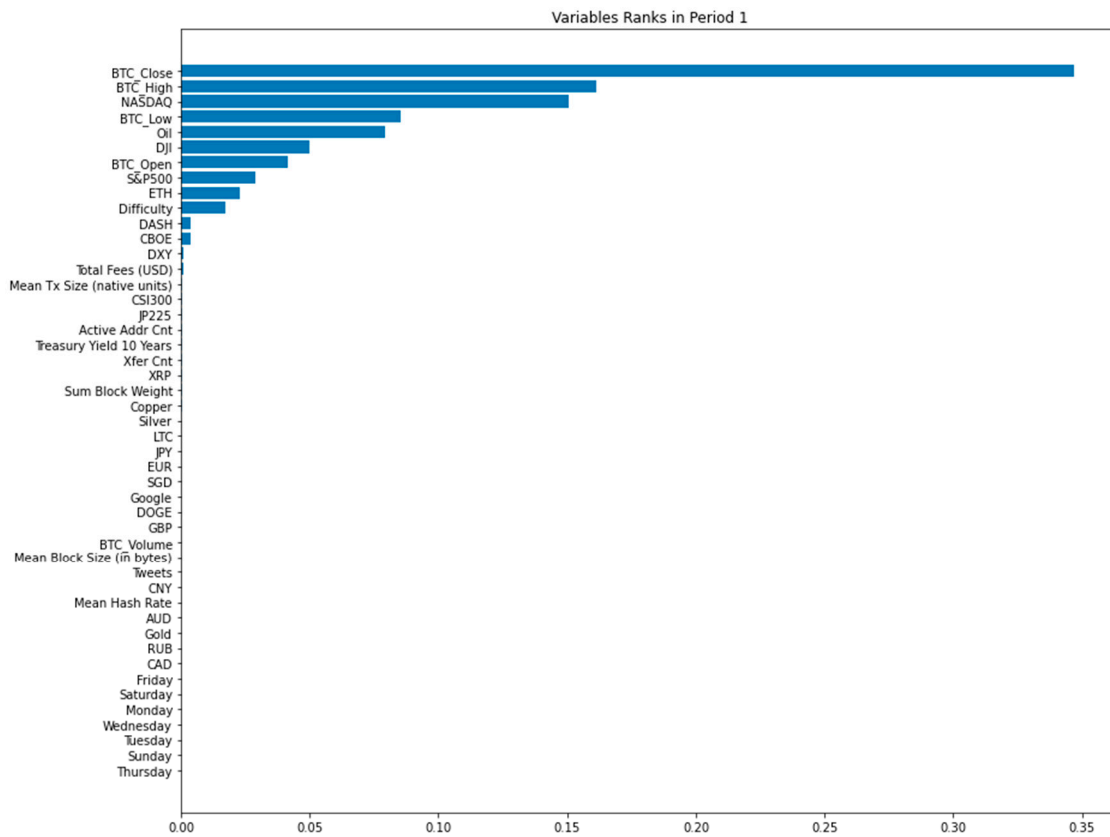


Figure 9. Cont.

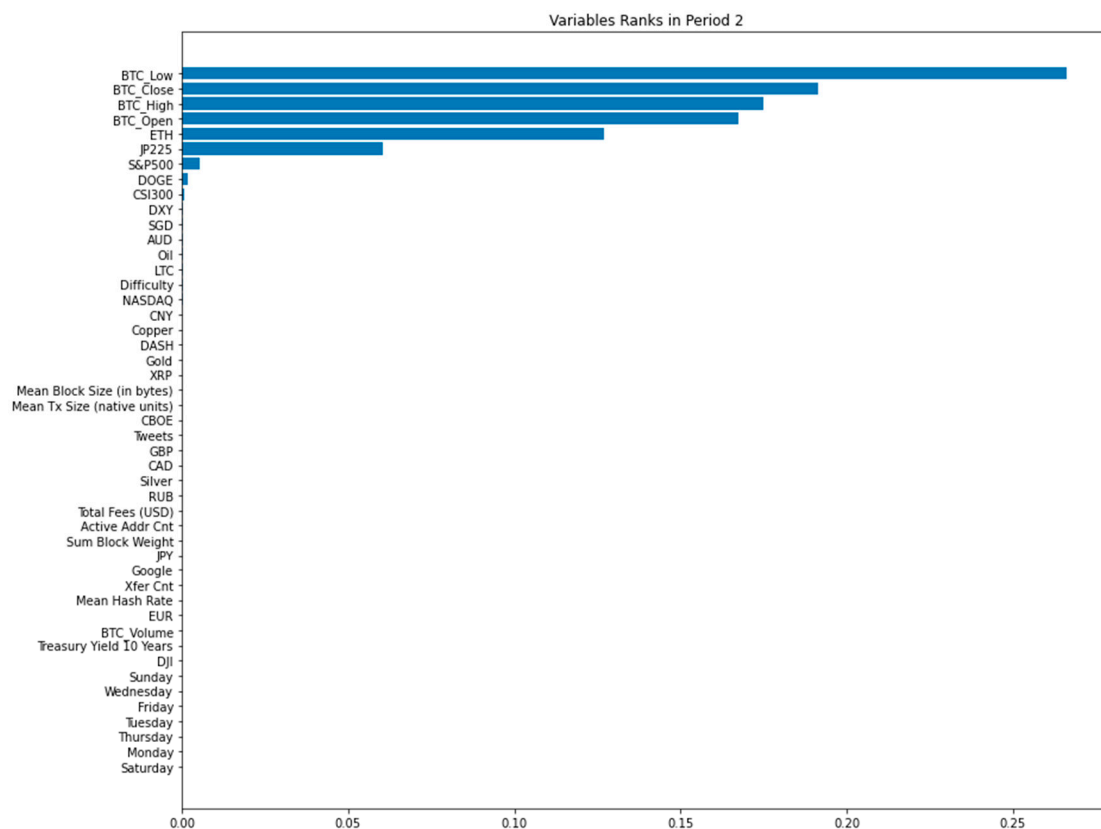


Figure 9. Explanatory variable importance ranks using random forest regression.

Whether it is Period 1 or Period 2, the importance of OHLC price of Bitcoin in the previous period is ranked high. However, what is interesting is that the relative order of open, high, low, and close in the two periods is not the same. According to the random walk theory of sequential prices, the price at each time point reflects the market’s expectation of the future price now, so the closing price closest in time should be the most important item among the four prices. The variable importance results for Period 1 accurately reflect this. However, in the ranking of Period 2, the lowest price of the previous period is considered the most important explanatory variable, and the closing price of the previous period is the last of these four prices. I think the possible reason that the lowest price in the previous period in Period 2 is important is related to the fact that there are more days of Bitcoin price decline in the later period of the Period 2 training sample, and the closing price is not at the highest level also implies that random forest regression delivers different results from random walk theory.

In addition to the variables of Bitcoin’s price, there are several other variables that are evaluated to be important when determining the price of Bitcoin. In Period 1, the NASDAQ index and crude oil prices in the United States are of high importance, even more important than the opening price of Bitcoin. From 7th to 10th places of importance are the American stock market index DJI, S&P500, ETH price, and the difficulty index of mining BTC. Among the top six explanatory variables other than Bitcoin price, the U.S. stock market index accounts for half of the three seats, which reflects the relationship between the U.S. stock market index and Bitcoin price from April 2015 to October 2018.

In Period 2, as shown in the Figure 9, since the importance of GBP in the 7th place is almost negligible, only the first six explanatory variables are considered. Except for the first four Bitcoin price variables, the remaining two are ETH, which is also a cryptocurrency, and Japan’s stock market index JP225.

Regarding the explanatory variables that determine the importance of Bitcoin prices, it can be summarized that the OHLC prices of Bitcoin itself in the previous period are the

most important. The importance of the remaining variables changes over time. The stock market index has the highest importance among all major categories. The feature of the high importance of the US stock market index in Period 1 has not been continued in Period 2. The importance of the Japanese stock market increased in Period 2. ETH is the only non-Bitcoin price variable that is considered important for Bitcoin price predictions in both Period 1 and Period 2.

In addition to obtaining the order of importance, to further study the impact of the presence or absence of explanatory variables on the prediction error, two additional tests were performed, which took turns taking out the least important and most important explanatory variable sets, respectively. The results are shown in Tables 5 and 6 below. The normal column is the importance ranking of all explanatory variables in Figure 9. The ascending column is to extract the most important explanatory variables and repeat the experiment. The descending column is to extract the least important explanatory variables and repeat the experiment. The results show that among the top variables in Period 1, except for BTC_Close, BTC_High, NASDAQ, and BTC_Low, all other variables have changed by more than two ranks. In contrast, the ranking of Period 2 is more stable, and the variables from 1 to 6 have not changed except for BTC_Open and BTC_Close.

Table 5. Summary of explanatory variables importance of Period 1.

Ranking (Period 1)	Normal	Ascending	Descending
1	BTC_Close	BTC_Close	BTC_Close
2	BTC_High	BTC_High	BTC_High
3	NASDAQ	NASDAQ	NASDAQ
4	Oil	BTC_Low	BTC_Low
5	BTC_Low	BTC_Open	BTC_Open
6	S&P500	Oil	DJI
7	BTC_Open	Difficulty	Oil
8	DJI	S&P500	S&P500
9	ETH	DJI	ETH
10	Difficulty	JP225	Difficulty

Table 6. Summary of explanatory variables importance of Period 2.

Ranking (Period 2)	Normal	Ascending	Descending
1	BTC_Low	BTC_Low	BTC_Low
2	BTC_High	BTC_High	BTC_High
3	BTC_Close	BTC_Close	BTC_Open
4	BTC_Open	BTC_Open	BTC_Close
5	ETH	ETH	ETH
6	JP225	JP225	JP225
7	S&P500	S&P500	CSI300
8	DOGE	DOGE	AUD
9	CSI300	GBP	NASDAQ
10	DXY	EUR	DXY

In addition to the ranking results, I analyzed the change in RMSE after taking out the most important variables in turn. The RMSE corresponding to the variable name on the abscissa refers to the RMSE error after removing it and the upper variables in Figure 10. Therefore, a large range of RMSE changes can show the importance of this variable relative to the remaining variables.

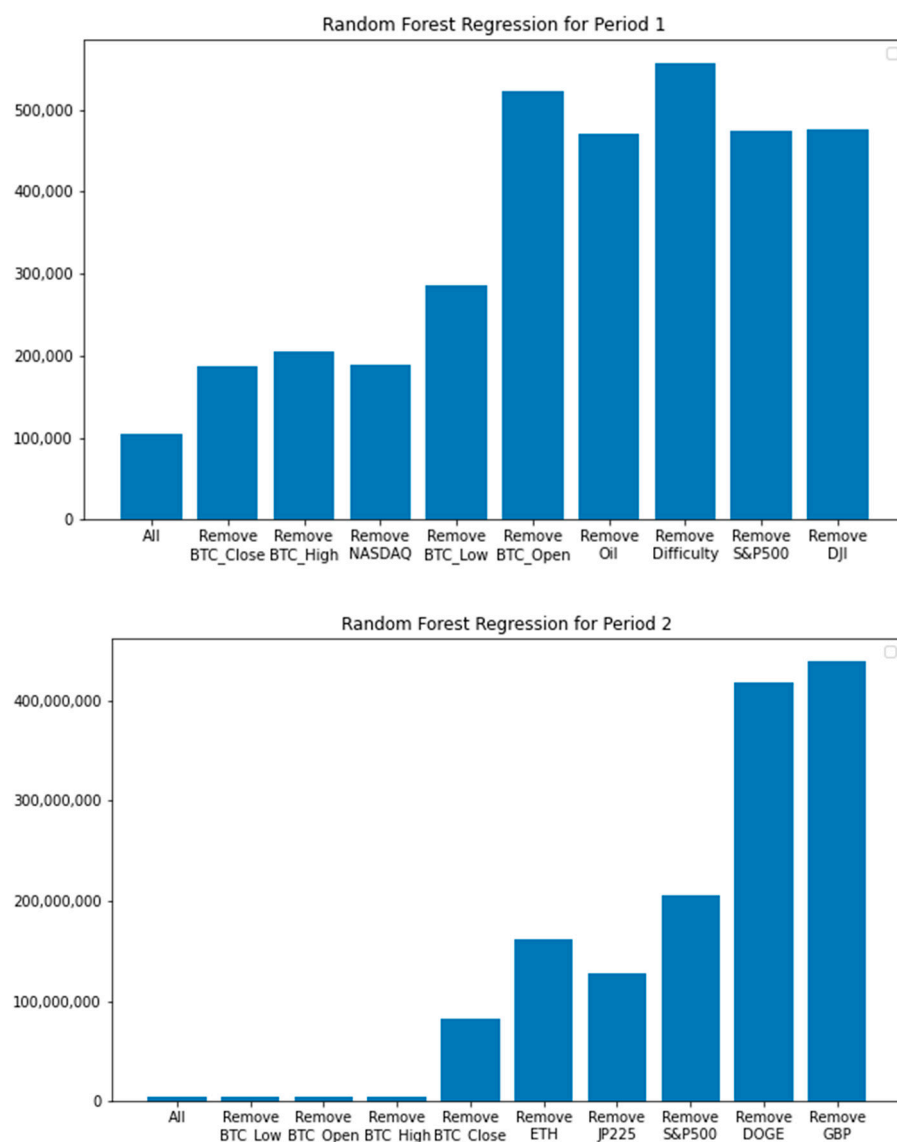


Figure 10. RMSE after removing the most important variable.

The significant increase in RMSE in Period 1 occurs when BTC_Open is removed. Removing BTC_Open corresponds to removing all OHLC from the previous day data, which shows that when predicting the next Bitcoin price, at least one of the current OHLC Bitcoin price needed. In Period 2, three large changes in RMSE occurred when BTC_Open, ETH, and DOGE were removed separately. Although the result of random forest shows that DOGE appears less often than JP225 and S&P500 in the nodes of all sub-regression trees, the sharp rise in RMSE after removing DOGE shows the effect of DOGE on prediction accuracy. The three large changes in Period 2 are all related to the price variables of cryptocurrency, indicating that the correlation between Bitcoin price and the cryptocurrency market has increased after 2018.

Based on the results about the importance of predicting the price of Bitcoin, I compared the prediction performance between the model with all variables and the model only with important variables (BTC_Close, BTC_High, NASDAQ, and BTC_Low for Period 1; BTC_Close, BTC_High, BTC_Low, BTC_Open, ETH, and JP225 for Period 2). The results show that the prediction accuracy of the model with all explanatory variables is better while the RMSE is 3% smaller than the results using only important variables (Figure 11).

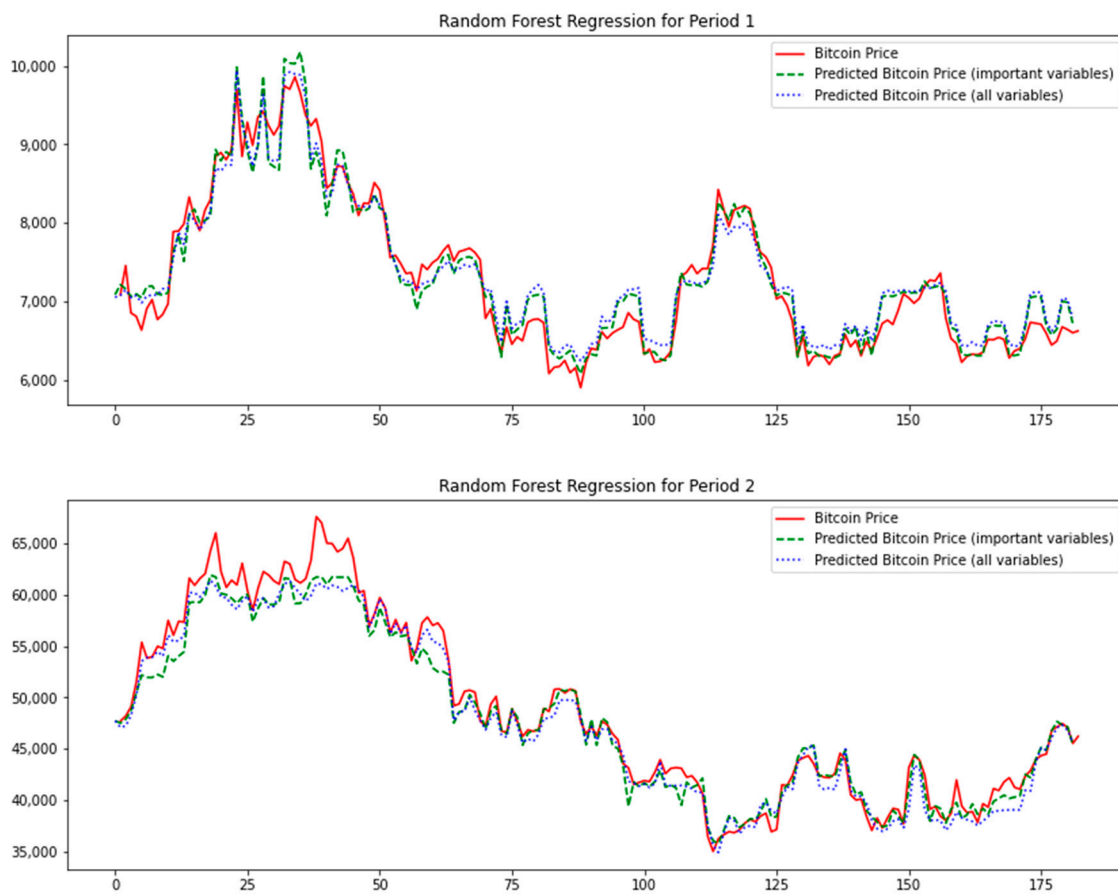


Figure 11. RFR results by all variables and only important variables.

5.2. Results of LSTM

I found that bringing redundant explanatory variables into the model for training leads to a decrease in model accuracy. The accuracy of the model obtained after all 47 explanatory variables are brought in is lower than that of the model using part of the variables, such as the lightweight model using only four Bitcoin price variables. On the contrary, if too few explanatory variables are used, the prediction accuracy of the model also reduces. For example, after adding some other variables to the lightweight model with four Bitcoin price OHLC variables, the prediction accuracy becomes better. Therefore, I have conducted a lot of experiments and attempts on what set of explanatory variables should be substituted in each period. Since there is no such problem in random forest due to it is ensemble algorithm, there is no need to discuss it in random forest regression.

Since the combination of explanatory variables brought in directly affects the prediction accuracy of the model, by referring to the importance rank of the explanatory variables using random forest regression, the respective explanatory variable sets of Period 1 and Period 2 are set in Table 7.

As the learning results of deep learning are related to the combination of randomly selected learning samples from the sample, randomness was present in the experimental results. Therefore, when comparing the model results, instead of comparing the accuracy of a single model, the average of the results of 30 experiments for various models is compared. The method of comparing the average of multiple experimental results was also applied in the experiments done by Liu et al. (2021).

Table 7. Explanatory variables used in Period 1 and Period 2.

	Period 1	Period 2
Variables	BTC_Open	BTC_Open
	BTC_High	BTC_High
	BTC_Low	BTC_Low
	BTC_Close	BTC_Close
	ETH	ETH
	Oil	JP225
	S&P500	
	NASDAQ	
	DJI	
	Difficulty	

The one-lagged accuracies of the models for two periods are shown in Table 8 and Figure 12.

Table 8. Errors of the LSTM models.

	Period 1	Period 2
RMSE	330.26	3045.87
MAPE	3.57%	4.68%

Note: the results are the average of 30 runs.

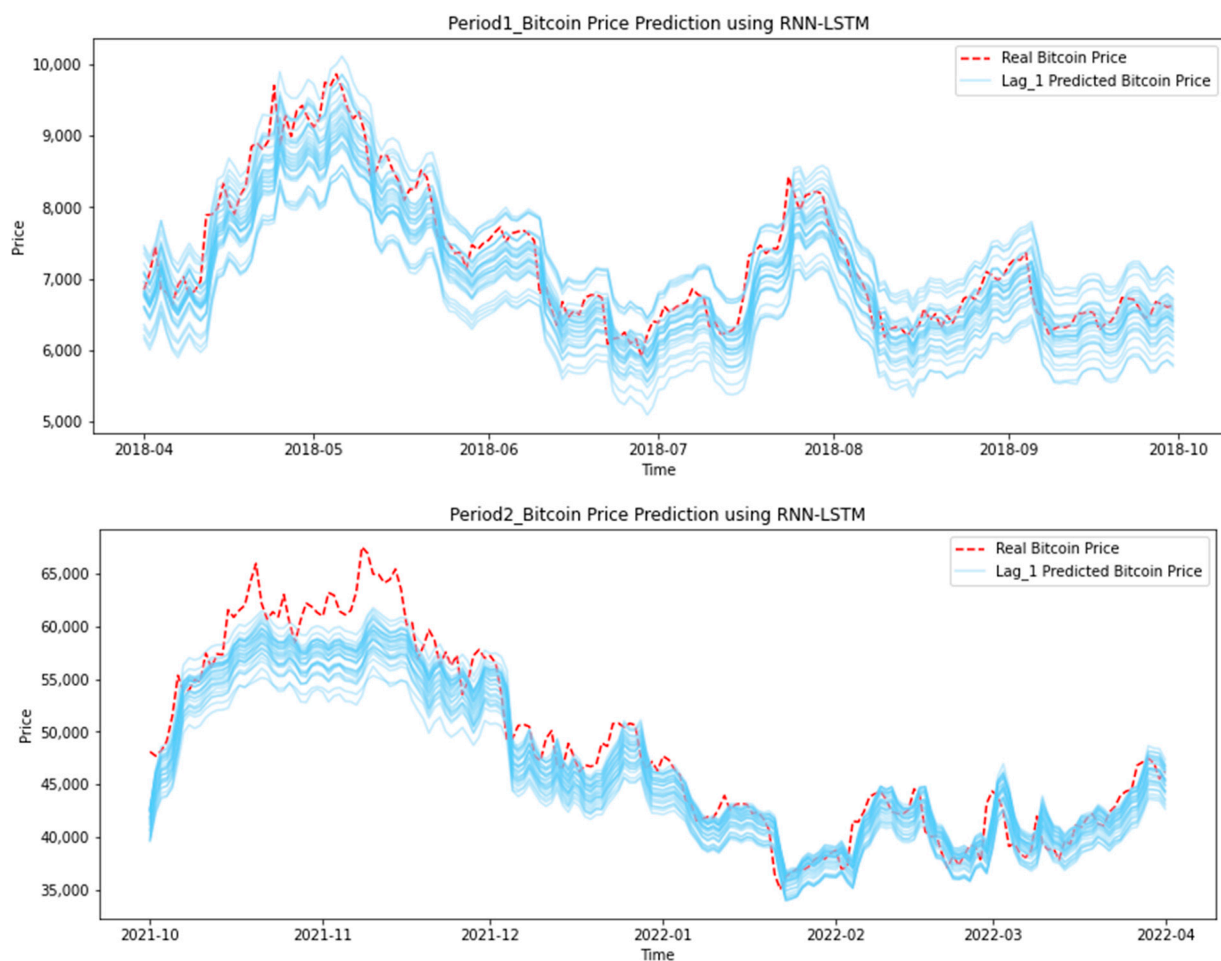


Figure 12. Comparison of the true price of Bitcoin and predicted price based on different models (LSTM).

When the Bitcoin price is greater than USD 60,000 in the early period of Period 2, the prediction results of the LSTM algorithm met the same problem of underrating as that in random forest regression. By comparing the MAPE of the two periods, the prediction accuracy of Bitcoin price in Period 1 is better than that in Period 2. This reflects that the correlation between Bitcoin and traditional markets has decreased in recent years, and the randomness of prices has increased. This result also echoes the conclusion that the price correlation is more and more determined by the previous period's own price, as reflected in the importance ranking of random forest regression in Period 2.

5.3. Relationship between Precision and Number of Variable Periods

Regarding the relationship between model accuracy and the number of lags of explanatory variables, I compared the results of five models with lags from 1 to 5. Whether it is Period 1 or Period 2, the conclusion is that the MAPE of random forest regression increase with the number of periods added as shown in Figure 13. Models trained by only explanatory variable data from the previous period had the best accuracy. This feature of the lagged relationship supports the efficient market hypothesis.

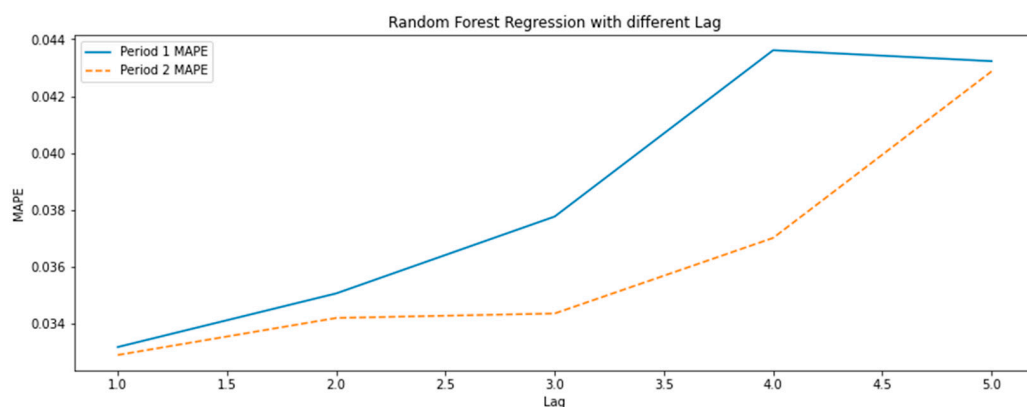


Figure 13. Relationship between MAPE and the number of lags (random forest regression).

LSTM is a deep learning algorithm with good predictive performance for time series data. The conclusion on whether it is necessary to refer to the data of multiple periods before when predicting the price of Bitcoin is that the prediction accuracy of the model that only needs the previous period is the best. As shown in the results of Period 1 (ten-variable model) and Period 2 (six-variable model) in Figure 14 below, although the price trends of each model are close to real price, the more periods of data substituted into the model, the smoother and smoother the curve of the forecast data becomes, deviating from the real price.

The conclusion on the number of data periods required when training the model is that only using the most recent period of data is sufficient. This conclusion is close to the efficient market hypothesis. The current price reflects the market's expectation of the future price of the asset, and the price of the previous period has no reference value.

According to summary Table 9, it can be found that whether it is Period 1 or Period 2, the price prediction accuracy of the random forest regression model with a lag of 1 is better than that of the LSTM algorithm.

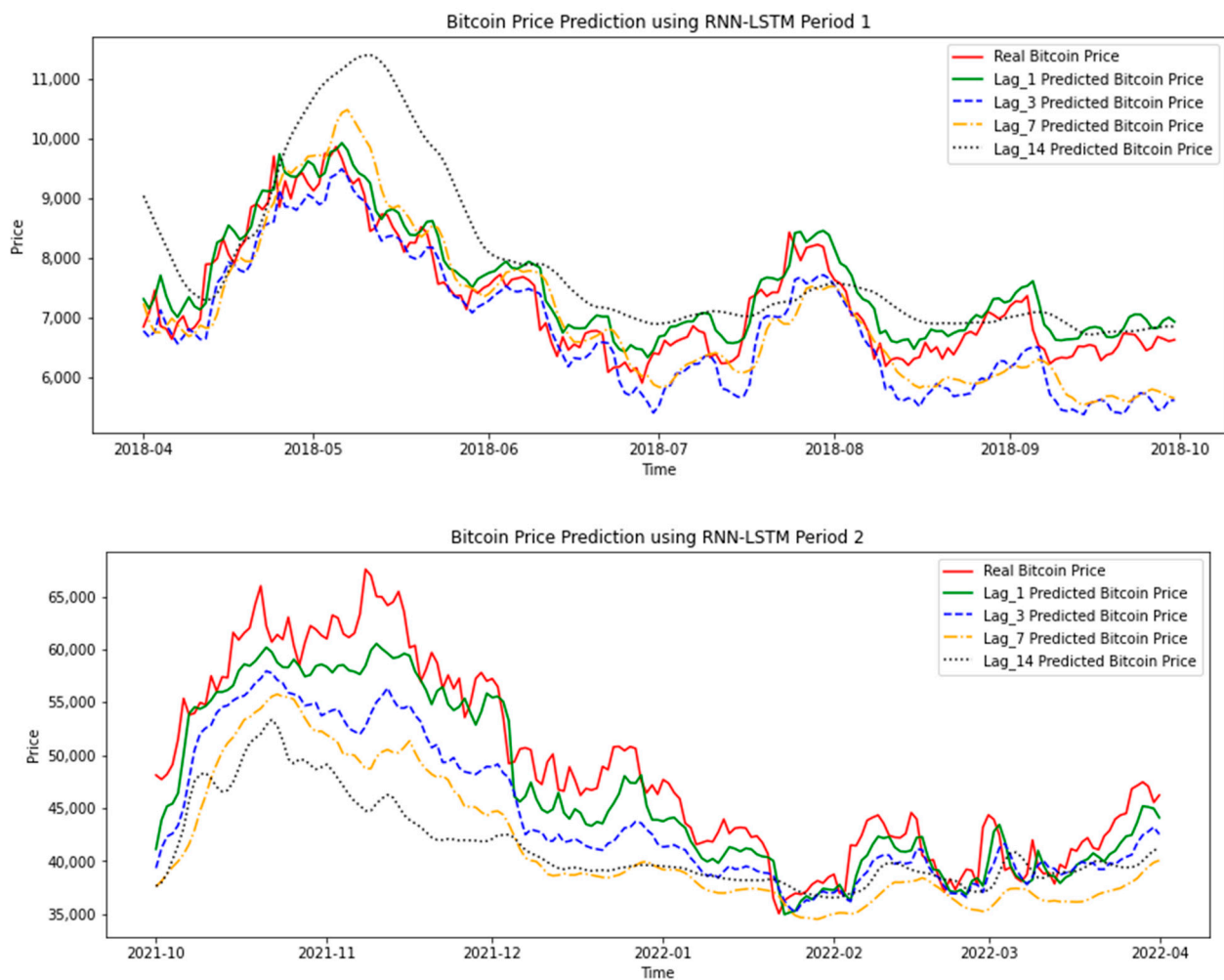


Figure 14. Relationship between accuracy and the number of lags (LSTM).

Table 9. Model evaluation using the accuracy and error.

		RMSE	MAPE	DA
Period 1	Random forest regression	321.61 (1.67%)	3.39%	51.93%
	LSTM	330.26 (1.71%)	3.57%	49.98%
Period 2	Random forest regression	2096.24 (3.48%)	3.29%	52.49%
	LSTM	3045.87 (5.05%)	4.68%	48.09%

Note: 1. The results are an average of 30 runs. 2. The model of LSTM is the Period 1 ten-variable model and the Period 2 six-variable model. 3. The brackets in the RMSE column are the values that have not been post-processed (min/max).

Except for the MAPE index of random forest regression, the other three groups (RMSE of random forest regression, RMSE and MAPE of LSTM) all showed that the prediction error of Period 2 is greater than that of Period 1. This result reflects that the Bitcoin price after October 2018 has become less predictable for the same algorithm. I think this result is related to the fact that the test data of Period 2 is in the bubble period, since machine learning is mainly based on the data of the training model when making predictions. Because the price of the bubble period is too high, the data of historical training are slightly

similar, causing the final accuracy declines. This phenomenon is obvious when random forest regression predicts the price of Bitcoin over USD 60,000.

Moreover, the comparison of error values does not reflect the results of hypothesis testing. I used the Diebold–Mariano test and the Clark–West test to further compare the significance of the prediction errors of the random forest regression and LSTM algorithms. The result is that no matter Period 1 or Period 2, the value is not greater than 1.64 required in the case of $\alpha = 95\%$. Thus, it cannot be denied that the prediction accuracy of LSTM is better than that of random forest regression.

Although random forest regression is not significantly better than LSTM, as an algorithm that has not been widely mentioned in the past literature, random forest regression has proven to be equivalent to or even better than LSTM in predicting the price of Bitcoin, as shown in Table 10.

Table 10. D–M test and C–W test results on the significant difference between random forest regression and LSTM.

	DM Test (MSE)	DM Test (MAE)	Clark and West Test
Period_1	0.36	0.52	0.84
Period_2	0.47	0.47	0.63

Note: when $\alpha = 95\%$, the statistical value of one-tailed test is 1.64.

6. Discussion

As a derivative comparison of experimental accuracy, Table 11 shows the DM Test and Clark–West test of random forest regression and LSTM relative to the prediction results of random walk. The results of the test show that the prediction accuracy of random walk is worse than that of random forest regression or LSTM and cannot be denied.

Table 11. D–M test and C–W test results on the significant difference between random walk and random forest regression or LSTM.

RFR/Random Walk	DM Test (MSE)	DM Test (MAE)	Clark and West Test
Period_1	0.33	0.39	0.58
Period_2	0.55	0.68	0.93
LSTM/random walk	DM test (MSE)	DM test (MAE)	Clark and West test
Period_1	0.34	0.41	0.47
Period_2	0.84	1.15	1.00

Note: when $\alpha = 95\%$, the statistical value of one-tailed test is 1.64.

There are two directions about future research, shortening the time interval of samples and automation. First, subject to the acquisition of historical data, the unit of the experimental sample this time is daily data, which leads to the prediction of the price has a problem of long interval. Moreover, within 24 h, the possibility of price forecast deviation due to unpredictable problems increases. To avoid the problems caused by the time units discussed above, in the future, I am going to collect the date with intervals of 1 h or 5 min only for the variables with high importance indicators in this experiment. Then, predictive analysis is performed on the new data through random forest regression and LSTM. The second direction of expansion is automation, which can be subdivided into automation of data acquisition and automation of prediction. Regarding the feasibility of Bitcoin predictions, Guarino et al. (2022) have conducted many experiments and believed that the high performance of neural networks in cryptocurrency prediction can be used for transactions. To obtain the predicted price provided by the model at any time, it is necessary to provide the latest data of explanatory variables to the model. A server can be set up on AWS (Amazon Web Services) to collect data prices of various trading websites in real time, and at the same time provide users with the future predicted price of Bitcoin

processed by LSTM and random forest regression in the form of an API interface. Moreover, the increase in the number of data collections can also solve the problem of long-time interval.

7. Conclusions

In this paper, to predict the price of Bitcoin on the next day, (a) Bitcoin price variables, (b) the specific technical features of Bitcoin, (c) other cryptocurrencies, (d) commodities, (e) market index, (f) foreign exchange, (g) public attention, and (h) dummy variables of the week, a total of eight categories (47 variables) were used as explanatory variables. Random forest regression has the better price prediction accuracy than LSTM. In previous research, LSTM was widely used and recognized as an algorithm with high accuracy when predicting Bitcoin prices. This paper uses the random forest regression machine learning algorithm, which has not been widely used by other researchers in the previous literature and obtains a result with higher prediction accuracy than LSTM. Although random forest regression has the disadvantage of being unable to predict the results that did not appear in the training samples. For example, when the price of Bitcoin broke the record high, random forest regression could not provide a higher price result than the previous historical high. But with the increase in Bitcoin transaction history, I think random forest regression will perform better when Bitcoin price stabilizes.

As a horizontal comparison with the research that also used daily as the time unit to predict Bitcoin, the RMSE error of random forest regression in this experiment (0.017 in Period 1 and 0.035 in Period 2) is better than is better than 0.045 of LSTM and 0.051 of GRU in Awoke et al.'s (2021) experiment, but worse than 0.009 for SDAE in Liu et al.'s (2021) experiment. I think it is difficult to compare prediction accuracy between different Bitcoin price prediction experiments. First, Bitcoin has many prices bubble periods, and whether the test data is in a bubble period has a great impact. For example, the RMSE error of random forest regression in Period 2 of this study is twice that of Period 1. Secondly, the samples of different unit time cannot be judged by the size of the test error. Interestingly, the models with the best accuracy in Awoke et al.'s (2021) experiments are the models with a lag of seven periods. This result is different from the conclusion in this paper that the optimal model only needs the latest explanatory variables.

The results of random forest regression also show the explanatory variables that determine the price of Bitcoin in various periods. In the first price bubble interval from April 2015 to October 2018, when predicting price on the next day, in addition to the price of the previous period of Bitcoin, the US stock market index (NASDAQ, DJI, and S&P500), the price of oil, ETH price, and the difficulty of finding blocks of Bitcoin, these six variables of mining difficulty also play an important role. During the second price bubble from October 2018 to April 2022, in addition to the OHLC prices of Bitcoin in the previous day, the price of ETH and Japan's JP225 index act a big role. When predicting the price of Bitcoin greater than USD 60,000 per coin at the end of 2021, random forest regression exposed the problem that it cannot predict values which is not in the training samples. However, the prediction accuracy for the price range below USD 60,000 is good.

In addition to the accuracy conclusion of a single model, the research results also found that whether it is random forest regression or LSTM algorithm, as the number of past periods of the substituted explanatory variables increases, the prediction accuracy of the model decreases. The model with the highest accuracy is the one that only substitutes explanatory variables in the past period. This conclusion is close to the classic efficient market hypothesis.

Funding: This research received no external funding.

Data Availability Statement: Data were obtained from <https://github.com/shiitake-github/jrfm-2156907-data> (accessed on 1 October 2022).

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Table A1. Definition of explanatory variables.

Variables	Description	Variables	Description
(a) Bitcoin		Oil	WTI crude oil price
BTC_Open	Bitcoin’s opening price	Treasury Yield 10 years	Treasury Yield 10 years
BTC_Close	Bitcoin’s closing price	(e) Market Index	
BTC_High	Bitcoin’s highest price of the day	S&P500	The Standard and Poor’s 500
BTC_Low	Bitcoin’s lowest price of the day	DJI	Dow Jones Industrial Average
BTC_Volume	Bitcoin transaction volume	CBOE	Chicago Board Options Exchange
(b) The specific technology features of Bitcoin		NASDAQ	National Association of Securities Dealers Automated Quotations
Active addr cnt	The sum count of unique addresses that were active in the network (either as a recipient or originator of a ledger change) on a given day.	JP225	The Nikkei 225
Xfer cnt	The sum count of transfers on a given day. Transfers represent movements of native units from one ledger entity to another distinct ledger entity. Only transfers that are the result of result from a transaction and(non-zero) value are counted.	CSI300	China Securities Index 300
Mean Tx size (native units)	The sum value of native units transferred is divided by the count of transfers (i.e., the mean size of a transfer) between distinct addresses at that interval.	(f) Foreign Exchange	
Total fees (USD)	The sum USD value of all fees paid by the user that makes the transactions on a given day. Fees do not include new issuance.	DEX	U.S. Dollar Index
Mean hash rate	The mean rate at which miners are solving hashes at a given rate. Hash rate is the speed at which computations are being completed across all miners in the network.	EUR	The number of Euros it takes to buy one dollar
Difficulty	The mean difficulty on a given day of finding a hash that meets the protocol-designated requirement (i.e., the difficulty of finding a new block).	GBP	The number of British pounds it takes to buy one dollar
Mean block size (in bytes)	The mean size (in bytes) of all blocks created on a given day.	JYP	The number of Japanese yen it takes to buy one dollar
Sum block weight	The sum count of blocks created that interval that was included in the main (base) chain on a given day.	CAD	The number of Canadian dollars it takes to buy one dollar

Table A1. Cont.

Variables	Description	Variables	Description
(c) Other cryptocurrencies		AUD	The number of Australian dollars it takes to buy one dollar
LTC	Price of one Litecoin in USD	SGD	The number of Singapore dollars it takes to buy one dollar
XRP	Price of one Ripple in USD	CNY	The number of Chinese yuan it takes to buy one dollar
DASH	Price of one Dash in USD	RUB	The number of Russian rubles it takes to buy one dollar
DOGE	Price of one Dogecoin in USD	(g) Public Attention	
ETH	Price of one Ethereum in USD	Google	Google Trend
(d) Commodities		Tweets	Number of daily Tweets
Gold	Gold price per ounce	(h) Week	
Silver	Silver price per ounce	Monday–Sunday	Dummy variable
Copper	Copper price per ounce		

References

- Aggarwal, Apoorva, Isha Gupta, Novesh Garg, and Anurag Goel. 2019. Deep Learning Approach to Determine the Impact of Socio Economic Factors on Bitcoin Price Prediction. Paper presented at 2019 Twelfth International Conference on Contemporary Computing (IC3), Noida, India, August 8–10.
- Akyildirim, Erdinc, Oguzhan Cepni, Shaen Corbet, and Gazi Salah Uddin. 2021. Forecasting mid-price movement of Bitcoin futures using machine learning. *Annals of Operations Research* 1–32. [CrossRef] [PubMed]
- Awoke, Temesgen, Minakhi Rout, Lipika Mohanty, and Suresh Chandra Satapathy. 2021. Bitcoin Price Prediction and Analysis Using Deep Learning Models. In *Communication Software and Networks*. Singapore: Springer, pp. 631–40.
- Basak, Suryoday, Saibal Kar, Snehanshu Saha, Luckyson Khaidem, and Sudeepa Roy Dey. 2019. Predicting the direction of stock market prices using tree-based classifiers. *The North American Journal of Economics and Finance* 47: 552–67. [CrossRef]
- Baur, Dirk G., and Lai Hoang. 2021. The Bitcoin gold correlation puzzle. *Journal of Behavioral and Experimental Finance* 32: 100561. [CrossRef]
- Baur, Dirk G., and Thomas Dimpfl. 2021. The volatility of Bitcoin and its role as a medium of exchange and a store of value. *Empirical Economics* 61: 2663–83. [CrossRef] [PubMed]
- Blake, R. 2019. An Econometric Analysis of the Relationship between Bitcoin & Gold. Available online: https://medium.com/@blake_richardson/an-econometric-analysis-of-the-relationship-between-bitcoin-gold-2018-584b4c63a17 (accessed on 21 September 2022).
- Carbó, José Manuel, and Sergio Gorjón. 2022. Application of Machine Learning Models and Interpretability Techniques to Identify the Determinants of the Price of Bitcoin. Banco de Espana Working Paper No. 2215. Available online: <https://ssrn.com/abstract=4087481> (accessed on 1 October 2022).
- Chen, Wei, Huilin Xu, Lifen Jia, and Ying Gao. 2020a. Machine learning model for Bitcoin exchange rate prediction using economic and technology determinants. *International Journal of Forecasting* 37: 28–43. [CrossRef]
- Chen, Yinghao, Xiaoliang Xie, Tianle Zhang, Jiaxian Bai, and Muzhou Hou. 2020b. A deep residual compensation extreme learning machine and applications. *Journal of Forecasting* 39: 986–99. [CrossRef]
- Derbentsev, Vasily, Natalia Datsenko, Vitalina Babenko, Olha Pushko, and Oleg Pursky. 2020. Forecasting Cryptocurrency Prices Using Ensembles-Based Machine Learning Approach. Paper presented at 2020 IEEE International Conference on Problems of Infocommunications. Science and Technology (PIC S&T), Kharkiv, Ukraine, October 6–9; pp. 707–12.
- Erdas, Mehmet Levent, and Abdullah Emre Caglar. 2018. Analysis of the relationships between Bitcoin and exchange rate, commodities and global indexes by asymmetric causality test. *Eastern Journal of European Studies* 9: 27–45.
- Fan, Liwei, Sijia Pan, Zimin Li, and Huiping Li. 2016. An ica-based support vector regression scheme for forecasting crude oil prices. *Technological Forecasting and Social Change* 112: 245–53. [CrossRef]
- García-Medina, Andrés, and Toan Luu Duc Huynh. 2021. What Drives Bitcoin? An Approach from Continuous Local Transfer Entropy and Deep Learning Classification Models. *Entropy* 23: 1582. [CrossRef]
- Guarino, Alfonso, Luca Grilli, Domenico Santoro, Francesco Messina, and Rocco Zaccagnino. 2022. To learn or not to learn? Evaluating autonomous, adaptive, automated traders in cryptocurrencies financial bubbles. *Neural Comput & Applic* 34: 20715–56. [CrossRef]

- Huang, Jia-Yen, and Jin-Hao Liu. 2020. Using social media mining technology to improve stock price forecast accuracy. *Journal of Forecasting* 39: 104–16. [CrossRef]
- Jagannath, Nishant, Tudor Barbulescu, Karam M. Sallam, Ibrahim Elgendi, Asuquo A. Okon, Braden McGrath, Abbas Jamalipour, and Kumudu Munasinghe. 2021. A Self-Adaptive Deep Learning-Based Algorithm for Predictive Analysis of Bitcoin Price. *IEEE Access* 9: 34054–66. [CrossRef]
- Jaquart, Patrick, David Dann, and Christof Weinhardt. 2021. Short-term bitcoin market prediction via machine learning. *The Journal of Finance and Data Science* 7: 45–66. [CrossRef]
- Khan, Wasiat, Mustansar Ali Ghazanfar, Muhammad Awais Azam, Amin Karami, Khaled H. Alyoubi, and Ahmed S. Alfakeeh. 2020. Stock market prediction using machine learning classifiers and social media, news. *Journal of Ambient Intelligence and Humanized Computing* 13: 3433–56. [CrossRef]
- Kim, Alisa, Y. Yang, Stefan Lessmann, Tiejun Ma, M.-C. Sung, and Johnnie E. V. Johnson. 2020a. Can deep learning predict risky retail investors? A case study in financial risk behavior forecasting. *European Journal of Operational Research* 283: 217–34. [CrossRef]
- Kim, Jong-Min, Seong-Tae Kim, and Sangjin Kim. 2020b. On the Relationship of Cryptocurrency Price with US Stock and Gold Price Using Copula Models. *Mathematics* 8: 1859. [CrossRef]
- Lamothe-Fernández, Prosper, David Alaminos, Prosper Lamothe-López, and Manuel A. Fernández-Gámez. 2020. Deep Learning Methods for Modeling Bitcoin Price. *Mathematics* 8: 1245. [CrossRef]
- Liu, Mingxi, Guowen Li, Jianping Li, Xiaoqian Zhu, and Yinhong Yao. 2021. Forecasting the price of Bitcoin using deep learning. *Finance Research Letters* 40: 101755. [CrossRef]
- Livieris, Ioannis E., Niki Kiriakidou, Stavros Stavroyiannis, and Panagiotis Pintelas. 2021. An Advanced CNN-LSTM Model for Cryptocurrency Forecasting. *Electronics* 10: 287. [CrossRef]
- Livieris, Ioannis E., Stavros Stavroyiannis, Emmanuel Pintelas, and Panagiotis Pintelas. 2020. A novel validation framework to enhance deep learning models intime-series forecasting. *Neural Computing and Applications* 32: 17149–67. [CrossRef]
- McNally, Sean, Jason Roche, and Simon Caton. 2018. Predicting the Price of Bitcoin Using Machine Learning. Paper presented at 26th Euromicro International Conference on Parallel, Distributed and Network-Based Processing (PDP), Cambridge, UK, March 21–23; pp. 339–43.
- Mudassir, Mohammed, Shada Bennbaia, Devrim Unal, and Mohammad Hammoudeh. 2020. Time-series forecasting of Bitcoin prices using high-dimensional features: A machine learning approach. *Neural Computing and Applications* 1–15. [CrossRef]
- Nakamoto, Satoshi. 2008. Bitcoin: A Peer-to-Peer Electronic Cash System. Available online: <https://bitcoin.org/bitcoin.pdf> (accessed on 7 October 2022).
- Parvez, Shaik Javed. 2022. Bitcoin price prediction using Random Forest Regression. *Journal of Positive School Psychology* 6: 4352–58.
- Phaladisailoed, Thearasak, and Thanisa Numnonda. 2018. Machine learning models comparison for bitcoin price prediction. Paper presented at 2018 10th International Conference on Information Technology and Electrical Engineering (ICITEE), Bali, Indonesia, July 24–26; pp. 506–11.
- Philip, Richard. 2020. Estimating permanent price impact via machine learning. *Journal of Econometrics* 215: 414–49. [CrossRef]
- Politis, Agis, Katerina Doka, and Nectarios Koziris. 2021. Ether price prediction using advanced deep learning models. Paper presented at 2021 IEEE International Conference on Blockchain and Cryptocurrency (ICBC), Sydney, Australia, May 3–6; pp. 1–3.
- Rizwan, Muhammad, Sanam Narejo, and Moazzam Javed. 2019. Bitcoin Price Prediction Using Deep Learning Algorithm. Paper presented at 2019 13th International Conference on Mathematics, Actuarial Science, Computer Science and Statistics (MACS), Karachi, Pakistan, December 14–15; pp. 56–60.
- Saadah, Siti, and A. A. Ahmad Whafa. 2020. Monitoring Financial Stability Based on Prediction of Cryptocurrencies Price Using Intelligent Algorithm. Paper presented at 2020 International Conference on Data Science and Its Applications (ICoDSA), Bandung, Indonesia, August 5–6; pp. 1–10.
- Sebastião, Helder, and Pedro Godinho. 2021. Forecasting and trading cryptocurrencies with machine learning under changing market conditions. *Financial Innovation* 7: 1–30. [CrossRef] [PubMed]
- Selmi, Refk, Walid Mensi, Shawkat Hammoudeh, and Jamal Bouoiyour. 2018. Is Bitcoin a hedge, a safe haven or a diversifier for oil price movements? A comparison with gold. *Energy Economics* 74: 787–801. [CrossRef]
- Shin, MyungJae, David Mohaisen, and Joongheon Kim. 2021. Bitcoin Price Forecasting via Ensemble-based LSTM Deep Learning Networks. Paper presented at 2021 International Conference on Information Networking (ICOIN), Jeju Island, Republic of Korea, January 13–16; pp. 603–8.
- Tandon, Sakshi, Shreya Tripathi, Pragya Saraswat, and Chetna Dabas. 2019. Bitcoin Price Forecasting using LSTM and 10-Fold Cross validation. Paper presented at 2019 International Conference on Signal Processing and Communication (ICSC), Noida, India, March 7–9; pp. 323–28. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

Article

Do Large Datasets or Hybrid Integrated Models Outperform Simple Ones in Predicting Commodity Prices and Foreign Exchange Rates?

Jin Shang and Shigeyuki Hamori * 

Graduate School of Economics, Kobe University, 2-1 Rokkodai, Nada-Ku, Kobe 657-8501, Japan;
susanfeir@yahoo.co.jp

* Correspondence: hamori@econ.kobe-u.ac.jp

Abstract: With the continuous advancement of machine learning and the increasing availability of internet-based information, there is a belief that these approaches and datasets enhance the accuracy of price prediction. However, this study aims to investigate the validity of this claim. The study examines the effectiveness of a large dataset and sophisticated methodologies in forecasting foreign exchange rates (FX) and commodity prices. Specifically, we employ sentiment analysis to construct a robust sentiment index and explore whether combining sentiment analysis with machine learning surpasses the performance of a large dataset when predicting FX and commodity prices. Additionally, we apply machine learning methodologies such as random forest (RF), eXtreme gradient boosting (XGB), and long short-term memory (LSTM), alongside the classical statistical model autoregressive integrated moving average (ARIMA), to forecast these prices and compare the models' performance. Based on the results, we propose novel methodologies that integrate wavelet transformation with classical ARIMA and machine learning techniques (seasonal-decomposition-ARIMA-LSTM, wavelet-ARIMA-LSTM, wavelet-ARIMA-RF, wavelet-ARIMA-XGB). We apply this analysis procedure to the commodity gold futures prices and the euro foreign exchange rates against the US dollar.

Keywords: hybrid forecasting approaches; two-step forecasting approaches; gold; euro; sentiment analysis; machine learning; ARIMA; wavelet transformation; seasonal decomposition; long short-term memory; random forest; eXtreme gradient boosting



Citation: Shang, Jin, and Shigeyuki Hamori. 2023. Do Large Datasets or Hybrid Integrated Models Outperform Simple Ones in Predicting Commodity Prices and Foreign Exchange Rates? *Journal of Risk and Financial Management* 16: 298. <https://doi.org/10.3390/jrfm16060298>

Academic Editor: Thanasis Stengos

Received: 27 April 2023

Revised: 5 June 2023

Accepted: 6 June 2023

Published: 9 June 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The increasing utilization of sentiment analysis (SA) for obtaining a sentiment index holds promise as an approach for predicting commodity prices and foreign exchange rates. By analyzing unstructured data such as social media posts, news articles, and other textual data, SA provides insights into public opinions and market sentiment, enabling price prediction (Smailović et al. 2013). Utilizing a sentiment index, rather than relying on a large dataset of indicators, offers several advantages, including simplifying the modeling process and reducing the risk of overfitting. SA also offers a more up-to-date perspective on market sentiment, as it captures real-time changes in public opinion and market sentiment (Philander and Zhong 2016). However, while a sentiment index proves valuable in predicting short-term fluctuations (Qiu et al. 2022) in commodity and foreign exchange markets, long-term trends in these markets are more significantly influenced by factors such as macroeconomic indicators and political events. Hence, while SA presents a promising approach to prediction, we must also consider its limitations and potential biases and supplement SA with other relevant data sources and indicators.

Meanwhile, research has demonstrated that advancements in machine learning and the availability of more data enhance the accuracy of price prediction in certain cases (Bakay and Ağbulut 2021; Bouktif et al. 2018; Wang and Wang 2016; Amat et al. 2018; Chatzis et al. 2018; Farsi et al. 2021; Zhang and Hamori 2020; Plakandaras et al. 2015; Luo et al. 2019;

McNally et al. 2018; Phyo et al. 2022; Nguyen and Ślepaczuk 2022). These technologies aid in identifying patterns and correlations within large and complex datasets that may prove challenging for human analysts to discern. However, employing large datasets and machine learning algorithms does not guarantee accuracy as these techniques are susceptible to biases, overfitting, and the appropriateness of the model design. In certain scenarios, simple models may outperform more sophisticated ones (He 2018), particularly when limited data are available or the underlying relationships are straightforward. Decision making and risk management may, at times, derive greater benefit from simple models based on relevant facts and hypotheses.

Recent research has garnered significant interest from academics and practitioners due to the emergence of hybrid techniques that combine classical models with machine learning models. Hybrid prediction models have been utilized in various research fields, including meteorology, hydraulics, and exhaust emissions, for forecasting purposes (Chang et al. 2019; Liu et al. 2018; de O. Santos Júnior et al. 2019; McNally et al. 2018; Sadefo Kamdem et al. 2020; Selvin et al. 2017; Xue et al. 2022; Sun et al. 2022; Wu et al. 2021; Wu and Wang 2022; Yu et al. 2020; Zhang et al. 2018, 2022; Zolfaghari and Gholami 2021; Ma et al. 2019; Dave et al. 2021; Zhao et al. 2022; Moustafa and Khodairy 2023; Zolfaghari and Gholami 2021). This study proposes several approaches that integrate machine and deep learning models with conventional statistical models, based on the assumption that time series can be decomposed into linear and nonlinear components or into time-dependent sums of frequency components and noise.

Hence, the primary objectives of this research are as follows: First, to analyze whether sentiment indicators derived from sentiment analysis techniques can outperform a large dataset of indicators when employing machine learning and deep learning methods for prediction. Second, to verify whether machine learning models, which have gained considerable attention, genuinely exhibit better prediction capabilities than classical ARIMA models. Third, to apply our proposed hybrid model to commodity gold futures prices and foreign exchange rates, evaluate their prediction performance, and compare them with the aforementioned machine learning and classical statistical approaches.

This study is divided into three steps. In the first step, we perform sentiment analysis on the collected unstructured news headlines to obtain a sentiment index (referred to as the SI dataset). Then, we calculate technical indicators and collect other relevant indicators from stock markets, bond markets, commodity markets, and foreign exchange markets to create a multivariate dataset (referred to as the large dataset). In the second step, we apply moving window machine learning approaches (RF, XGB, and LSTM) and a classical statistical model (ARIMA) to these two datasets to evaluate their prediction performance using the root mean squared error (RMSE), mean absolute percentage error (MAPE), and mean absolute error (MAE). In the third step, we propose several decompositions and transformations integrated with statistical and machine learning approaches, such as seasonal-decomposition-ARIMA-LSTM, wavelet-ARIMA-LSTM, wavelet-ARIMA-RF, and wavelet-ARIMA-XGB. Specifically, we first transform and decompose the time series into linear and nonlinear parts or dynamic levels and noise parts. Then, we apply classical ARIMA to predict the linear and dynamic levels and use RF, XGB, and LSTM machine/deep learning approaches to predict the nonlinear and noise parts. We evaluate our proposed approaches using RMSE, MAPE, and MAE and compare the prediction results with the aforementioned forecasting. Additionally, we perform walk-forward testing to validate the effectiveness of the triple-combination approaches. To assess any statistically significant differences between our proposed approach and the ARIMA model, we utilize the modified Diebold–Mariano test statistic. This comprehensive testing methodology provides further insights into the performance and comparative analysis of the proposed approaches.

The main findings of this study are as follows: First, the combination of the sentiment indicator with the moving window LSTM machine learning model demonstrates outstanding forecasting performance. Second, the sentiment indicator dataset used in conjunction with the moving window machine learning and deep learning models does not surpass

the performance of the traditional ARIMA model. Third, our proposed triple-combination approaches exhibit superior prediction performance compared to either the machine learning models or the ARIMA model when forecasting commodity gold futures prices and euro foreign exchange rates. Lastly, although the sentiment indicator dataset does not outperform the prediction accuracy of the ARIMA model, our empirical results indicate that the sentiment dataset is more accurate in predicting commodity prices and foreign exchange rates than the large dataset, which comprises various indicators.

To the best of our knowledge, this study is the first to investigate whether sentiment indicators can replace a large dataset of indicators in forecasting commodity prices and foreign exchange rates. Moreover, this study introduces a novel approach by combining data decomposition with machine learning models and classical statistical models to predict prices in commodity and foreign exchange markets. Additionally, the proposed triple-combination approaches demonstrate higher accuracy compared to the individual models. These findings offer new insights and potential predictors for investors and policymakers.

The rest of this paper is organized as follows: Section 2 reviews the literature. Section 3 provides a detailed description of the study's data, methodologies, and evaluation measures. Section 4 presents and analyzes the empirical results. Finally, Section 5 concludes the study.

2. Literature Review

A wide range of valuable Internet data, particularly textual data such as news press releases, are being evaluated for forecasting purposes in various fields, thanks to the rapid expansion of the Internet and advancements in big data technologies. Consequently, researchers are actively working on improving sentiment analysis (SA) predictions and exploring the potential of SA to enhance time series forecasting performance in different markets (Bollen et al. 2011; Naeem et al. 2021; Deeney et al. 2015; Li et al. 2016; Das et al. 2018; Pai and Liu 2018; Razzaq et al. 2019; Bedi and Khurana 2019; Ito et al. 2019, 2020; Sivri et al. 2022; Seals and Price 2020; Xiang et al. 2021; Guo et al. 2020; Sharma et al. 2020; Mukta et al. 2022).

The contribution of Bedi and Khurana (2019) is focused on improving SA prediction for textual data by incorporating fuzziness with deep learning. Ito et al. (2019) and Ito et al. (2020) propose a novel neural network model called the contextual sentiment neural network (CSNN) model, which offers insights into the SA prediction process and utilizes an initialization propagation (IP) learning strategy. Leveraging SA on Twitter tweets, Naeem et al. (2021) suggest a machine learning-based strategy for forecasting exchange rates. Their findings demonstrate that SA can facilitate the prediction of foreign exchange rates, particularly the US dollar against the Pakistani rupee. Li et al. (2016) acknowledge the usefulness of online data, including news releases and social media networks such as Twitter, in forecasting price changes. Xiang et al. (2021) propose a Chinese Weibo SA algorithm that combines the BERT (Bidirectional Encoder Representations from Transformers) model and the Hawkes process to effectively monitor changes in users' emotional states and perform SA on Weibo. However, limited studies have examined whether sentiment indicators can replace large sets of index data for forex prediction. If sentiment indicators can effectively replace a substantial amount of index datasets and achieve comparable or better forecasting performance, it could significantly enhance forecasting efficiency and provide valuable insights to investors and decision-makers.

Moreover, in recently published research, the use of rapidly developing machine and deep learning modeling techniques for forecasting time series is one of the most extensively researched topics in the academic literature (Bakay and Ağbulut 2021; Bouktif et al. 2018; Wang and Wang 2016; Amat et al. 2018; Chatzis et al. 2018; Farsi et al. 2021; Zhang and Hamori 2020; Plakandaras et al. 2015; Luo et al. 2019; McNally et al. 2018; Phyo et al. 2022). Specifically, Amat et al. (2018) demonstrate that fundamentals from simple exchange rate models (such as purchasing power parity (PPP) or uncovered interest rate parity (UIRP)) or Taylor-rule-based models improve exchange rate forecasts for major currencies when using machine learning models. Similarly, Zhang and Hamori (2020) find that integrating machine

learning models with traditional foreign exchange rate models and Taylor's rule foreign exchange rate models effectively predict foreign exchange rates. Phyo et al. (2022) train five of the best ML algorithms, including the extra trees regressor (ETR), random forest regressor (RFR), light gradient boosting machine (LGBM), gradient boosting regressor (GBR), and K neighbors regressor (KNN), to build the proposed voting regressor (VR) model. Li et al. (2020) propose a new dynamic ensemble forecasting system based on a multi-objective intelligent optimization algorithm to forecast the air quality index, which includes time-varying parameter weights and three main modules: a data preprocessing module, a dynamic integration forecasting module, and a system evaluation module. Plakandaras et al. (2015) predict daily and monthly exchange rates using machine learning techniques. Building on these empirical results, this paper considers the application of machine learning and deep learning methodologies to investigate whether sentiment indicator datasets can substitute for large datasets.

On the other hand, as a classical statistical model, ARIMA is used for long-term prediction (Darley et al. 2021). Many studies compare ARIMA and machine learning in forecasting time series (Shih and Rajendran 2019; Siami-Namini et al. 2018, 2019; He 2018; Yamak et al. 2019; Ribeiro et al. 2020; Liu et al. 2021). Siami-Namini et al. (2018) compare the ARIMA model with the LSTM model in forecasting time series and demonstrate that deep learning approaches such as LSTM outperform traditional models such as ARIMA. In contrast, He (2018) explores weekly crude oil price data from the U.S. Energy Information Administration between 2009 and 2017 to test the forecasting accuracy of time series models (simple exponential smoothing (SES), moving average (MA), and autoregressive integrated moving average (ARIMA)) against machine learning support vector regression (SVR) models. The main contribution of this study is to determine whether ARIMA provides more accurate forecasting results for crude oil prices than SVR models. Siami-Namini et al. (2019) conduct a behavioral analysis and comparison of BiLSTM and LSTM models and compare the two models with the ARIMA model. The results demonstrate that BiLSTM models provide better predictions compared to ARIMA and LSTM models. Yamak et al. (2019) conduct a comparison analysis between ARIMA, LSTM, and gated recurrent unit (GRU) for time series forecasting. Ribeiro et al. (2020) compare two benchmarks (autoregressive integrated moving average (ARIMA) and an existing manual technique used at the case site) against three deep learning models (simple recurrent neural networks (RNN), long short-term memory (LSTM), and gated recurrent unit (GRU)) and two machine learning models (support vector regression (SVR) and random forest (RF)) for short-term load forecasting (STLF) using data from a Brazilian thermoplastic resin manufacturing plant. Their empirical results show that the GRU model outperforms all other models. Liu et al. (2021) propose a seasonal autoregressive integrated moving average (SARIMA) model to predict hourly measured wind speeds in the coastal and offshore areas of Scotland. Motivated by the results of the prior literature and considering the limited literature comparing ARIMA models with machine learning and deep learning models for predicting gold prices and Euro FX prices, this study aims to fill this gap in the literature.

Since we are unable to demonstrate that machine learning and deep learning techniques outperform the traditional ARIMA model, we aim to enhance the accuracy of commodity price and foreign exchange rate predictions. In our literature research, we discover numerous studies in various fields, such as astronomy, hydraulics, exhaust emissions, and meteorology, that employ a combination of traditional models and other techniques such as machine learning, deep learning methodologies, and two-step models, which involve preprocessing the data before predicting time series. Some relevant studies include Chang et al. (2019), Liu et al. (2018), de O. Santos Júnior et al. (2019), McNally et al. (2018), Sadefo Sadefo Kamdem et al. (2020), Selvin et al. (2017), Xue et al. (2022), Sun et al. (2022), Wu et al. (2021), Wu and Wang (2022), Yu et al. (2020), Zhang et al. (2018, 2022), Zolfaghari and Gholami (2021), Ma et al. (2019), Dave et al. (2021), Zhao et al. (2022), Moustafa and Khodairy (2023), and Zolfaghari and Gholami (2021).

To enhance prognostic accuracy, Ma et al. (2019) propose a data-fusion approach that combines long short-term memory (LSTM), recurrent neural network (RNN), and the autoregressive integrated moving average (ARIMA) method to forecast fuel cell performance. Chang et al. (2019) present an electricity price-prediction model based on a hybrid of the LSTM neural network and wavelet transform. Liu et al. (2018) attempt to forecast wind speed using a deep learning strategy with wavelet transform. Dave et al. (2021) aim to provide accurate predictions of Indonesia's future exports by developing an integrated machine learning model with ARIMA. Zhou et al. (2022) propose a combined model based on complete ensemble empirical mode decomposition with adaptive noise (CEEMDAN), four deep learning (DL) models, and the autoregressive integrated moving average (ARIMA) model. Zhao et al. (2022) address the lack of using coupled models to separately model different frequency subseries of precipitation series for prediction and propose a coupled model based on ensemble empirical mode decomposition (EEMD), long short-term memory neural network (LSTM), and autoregressive integrated moving average (ARIMA) for month-by-month precipitation prediction. Moustafa and Khodairy (2023) implement four models, including long short-term memory (LSTM), autoregressive integrated moving average (ARIMA), seasonal autoregressive integrated moving average (SARIMA), and a hybrid model, to forecast the maximum sunspot number of cycles 25 and 26. Zolfaghari and Gholami (2021) employ a hybrid model that combines adaptive wavelet transform (AWT), long short-term memory (LSTM), and models from the ARIMAX-GARCH family to forecast stock indices for the Dow Jones Industrial Average (DJIA) and the Nasdaq Composite (IXIC). Chen and Wang (2019) integrate the LSTM and ARIMA models for predicting satellite time series data. Inspired by these studies, this investigation aims to propose hybrid approaches applicable to time series forecasting in commodity markets and foreign exchange markets.

To summarize, researchers have dedicated significant efforts to enhancing the accuracy of price prediction by utilizing machine learning techniques and internet-based information. The increasing availability of data sources, particularly textual data such as news articles, and advancements in big data technologies have led to the evaluation of various datasets for forecasting in different domains. However, in the context of time series forecasting in commodity and foreign exchange markets, there is a lack of literature that thoroughly compares the effectiveness of sentiment indicator datasets with large datasets containing diverse variables. Additionally, the recent academic literature extensively explores the application of rapidly evolving machine learning and deep learning modeling techniques for time series forecasting. Nevertheless, further investigation is required to determine whether machine learning and deep learning models outperform classical statistical methods, such as the ARIMA model, which have long been used for forecasting purposes in the commodity and foreign exchange markets. Therefore, in our study, we focus on improving forecasting accuracy by combining traditional models with other methods, including machine learning, deep learning techniques, and two-step models. We draw inspiration from previous studies conducted in fields such as astronomy, hydraulics, exhaust emissions, and meteorology, which have employed time series forecasting in their respective domains.

3. Data and Methodology

3.1. Data

3.1.1. Data Collection

Gold prices are widely regarded as a leading indicator of economic conditions, particularly inflation and market volatility, making it an extremely important commodity (Blose 2010; Livieris et al. 2020). As a result, gold is a popular investment asset (Ratner and Klein 2008) and is commonly used as a hedge against inflation and market volatility (Chua and Woodward 1982). Predicting gold prices can provide valuable insights for economic forecasts and assist policymakers and investors in making informed decisions (Raza et al. 2018). Additionally, many central banks maintain gold reserves as a means of preserving value and protecting against currency fluctuations (Aizenman and Inoue 2013).

On the other hand, foreign exchange rates have been utilized as leading indicators of economic growth and inflation (Razzaque et al. 2017). The foreign exchange market plays a crucial role in international trade (Latief and Lefen 2018), financial instrument settlement, inflation control, and overall economic development and currency stability. Accurate predictions of foreign exchange rates are essential for businesses and investors to develop effective hedging strategies that mitigate risks associated with currency fluctuations. Moreover, such predictions inform government policy decisions related to trade, monetary policy, and capital flows (Amato et al. 2005; Mussa 1976). Governments can use exchange rate predictions to anticipate the impacts of policy decisions on the economy and make necessary adjustments. It is also worth noting that the euro is the second-most traded currency globally, following the US dollar, and is extensively used by numerous European Union members. Given the widespread usage of the euro in international trade and its status as a major reserve currency, exchange rate fluctuations can significantly influence the costs and risks associated with international transactions. Therefore, forecasting euro exchange rates is vital for financial stability and effective hedging strategies. Consequently, this study selected gold futures prices from the commodity market and the EUR foreign exchange rate as the objects of forecasting.

Based on the concept of proposing a powerful alternative sentiment indicator to replace large datasets, this study applies sentiment analysis to unstructured data extracted from news headlines. The prediction objects selected for this study are gold futures prices and the euro exchange rate against the US dollar, sourced from invest.com. After preprocessing the dataset, a total of 3957 daily data points were obtained, covering the period from 3 February 2004 to December 2019. The prediction conducted in this study is one-day-ahead forecasting.

The large dataset used in this study consists of 22 different financial indicators obtained from various sources such as Bloomberg, Thomson Reuters Datastream, the Federal Reserve Bank, Investing.com, Yahoo! Finance, and Macrotrends. Specifically, the large dataset includes the stock market index, 10-year government bond yields, volatility indices, and significant commodity market indices such as oil, gas, corn, and wheat. Additionally, it incorporates 10 calculated technical indices, including moving averages, exponential weighted moving averages, Bollinger bands, moving average convergence divergence, and the relative strength index.

3.1.2. Sentiment Analysis and Sentiment Indicator

In this study, we conduct sentiment analysis to obtain a sentiment indicator as an input variable.

First, we utilize unstructured daily news headline text data from 19 February 2003 to 31 December 2020. The data consist of 1,226,258 news headlines collected from a reputable news source, the Australian Broadcasting Corporation (ABC). The news headline data are sourced from Harvard Dataverse, which was created by Kulkarni (2018). According to the authors' notes, "with a volume of two hundred articles each day and a good focus on international news, we can be fairly certain that every event of significance has been captured here".

For sentiment analysis on daily news headlines, we employ a Python natural language processing library called TextBlob. TextBlob is chosen for its ability to provide rules-based sentiment scores and assign polarity and subjectivity to words and phrases. These scores are derived from a pre-defined set of categorized words readily available from the Natural Language Toolkit (NLTK) database (Vijayarani and Janani 2016). The input data for sentiment analysis typically consist of a corpus, such as a collection of text documents. The output of sentiment analysis includes a sentiment polarity score (indicating positivity or negativity) and a subjectivity score (measuring opinionated-ness). The polarity score ranges from -1.0 to 1.0 , where -1.0 represents strong negativity and 1.0 represents high positivity. The subjectivity score ranges from 0.0 to 1.0 , where 0.0 denotes extreme objectivity or factual content, while 1.0 signifies high subjectivity.

The sentiment analysis procedure is described as follows:

- Firstly, the NLTK is used to clean the unstructured text data.
- Secondly, TextBlob is applied to classify the polarity and subjectivity of each news headline.
- Thirdly, the total number of subjective, objective, negative, positive, and neutral news headlines is counted for each day, and then divided by the total number of news headlines on that day.
- Fourthly, the sentiment analysis output data are obtained, which includes the percentage values for subjectivity, objectivity, negativity, neutrality, and positivity for each day.
- Finally, following Henry’s finance-specific dictionary (Henry and Leone 2016), the sentiment can be evaluated using the formula below:

$$SI_t = \frac{N_p(H_t) - N_n(H_t)}{N_p(H_t) + N_n(H_t)} \quad (1)$$

where H_t represents the collected news article headlines at time t , N_p represents the total number of positive news headlines in H_t , N_n represents the total number of negative news headlines in H_t , and SI_t represents the corresponding sentiment indicator.

The sentiment indicator represents the percentage difference between the number of positive and negative news articles.

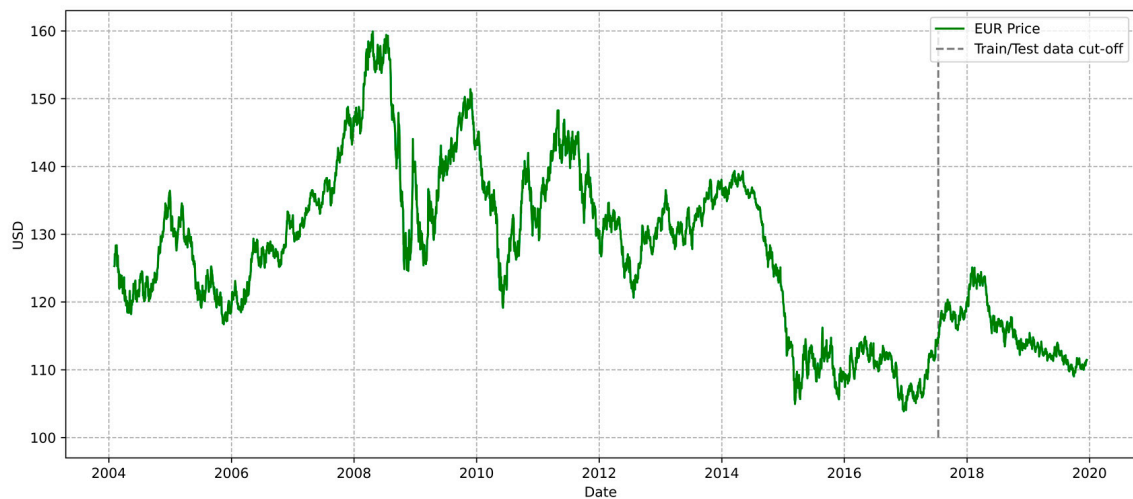
3.1.3. Sentiment Indicator Dataset and Large Indicator Dataset

After data processing, we obtain 3957 daily data points that contain 32 explanatory variables, covering a 15-year period from 3 February 2004 to 16 December 2019. The descriptions and sources of the data are elaborated in Table A1 of the Appendix A.

In this study, we use 85% of the daily data (3363 days) to train various models based on RF, XGBoost, and LSTM models. We then validate the remaining data (594 days) to conduct out-of-sample forecasting. Figure 1 illustrates the raw data of the gold futures prices, Figure 2 presents the prices of the euro rates multiplied by 100, and Figure 3 presents the calculated sentiment index based on the results of sentiment analysis. The dashed vertical line (14 July 2017) denotes the separation between the training and test data.

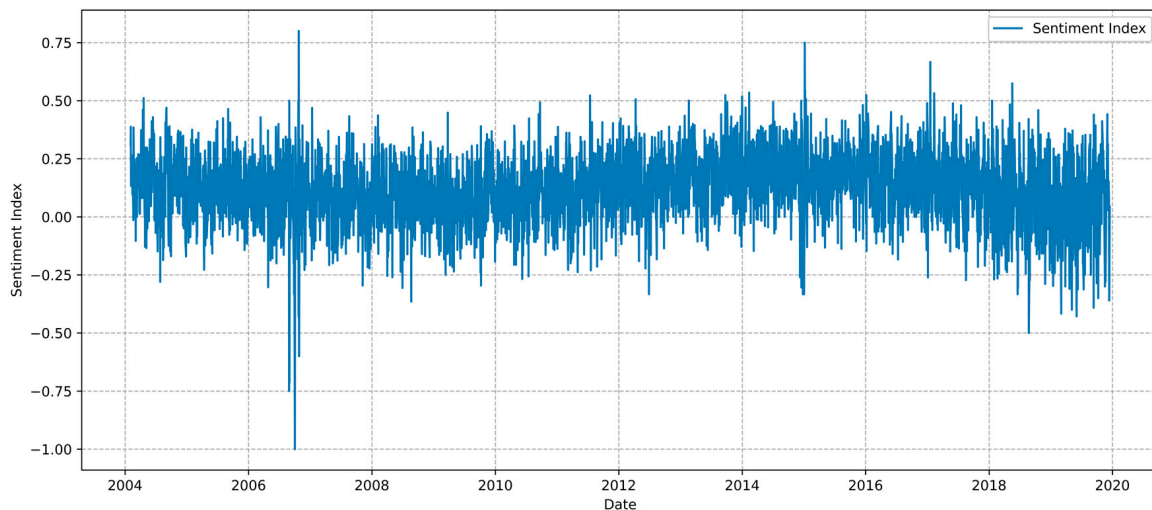


Figure 1. Historical data plotting for gold futures price. Note: This figure illustrates the raw data of gold futures prices and the dotted line represents the train/test data.



Euro Price

Figure 2. Historical data plotting for euro price. Note: This figure illustrates the raw data of the euro exchange rate multiplied by 100, and the dotted line represents the train/test data.



Sentiment Index

Figure 3. Sentiment index plotting. Note: This figure illustrates the calculated sentiment index based on the results of sentiment analysis.

To test the hypothesis that the sentiment indicator can be a substitute for the large datasets of indicators in exchange rate prediction, we construct two datasets to evaluate the effectiveness of the sentiment indicator and compare their predictive performance. Detailed information regarding these variables is provided in Table 1.

Table 1. Datasets used to predict gold futures prices and the euro exchange rates.

	Containing Variables	Number of Variables
SI dataset	Today's price + Sentiment Indicator	2
Large dataset	Today's price + Collected/Calculated Indicators + Sentiment Indicator	33

Note: SI dataset represents the dataset comprising of today's price and sentiment indicator. Large dataset represents the dataset comprising of today's price, sentiment indicator, and collected/calculated indicators.

3.2. Prediction Models and Proposed Approaches

This study applies the RF, XGB, and LSTM approaches in combination with the expanding moving window (EMW), and fixed moving window (FMW) methods to predict gold futures commodity prices and the euro foreign exchange rate. The initial parameters (Wysocki and Ślepaczuk 2022) are selected using the grid search method. Specifically, trained models with time-varying parameters are used to predict one-period-ahead prices, and the prediction performance of these models is evaluated using the remaining test datasets. The moving window technique proceeds iteratively with the prediction, where the size of the expanding moving window or fixed moving window is extended or shifted by one-time step in each iteration. Furthermore, the study employs the widely applied time series forecasting model ARIMA to validate the superiority of the sentiment indicator dataset. Additionally, triple-combination approaches are proposed, including wavelet-ARIMA-LSTM (wavelet-ARIMA-RF/wavelet-ARIMA-XGB) and seasonal-decomposition-ARIMA-LSTM.

3.2.1. Expanding Moving Window (EMW) and Fixed Moving Window (FMW)

This study employs two patterns of moving window techniques to predict one-period-ahead, aiming to investigate whether there is a difference in prediction performance when excluding historical data. One pattern is the fixed-length moving window (FMW) technique, and the other is the expanding-length moving window (EMW) technique.

The moving window statistics proceed iteratively with the prediction, extending or shifting the size of EMW or FMW by one time step in each iteration. Figure 4 illustrates the mechanism of EMW, while Figure 5 depicts the mechanism of FMW.

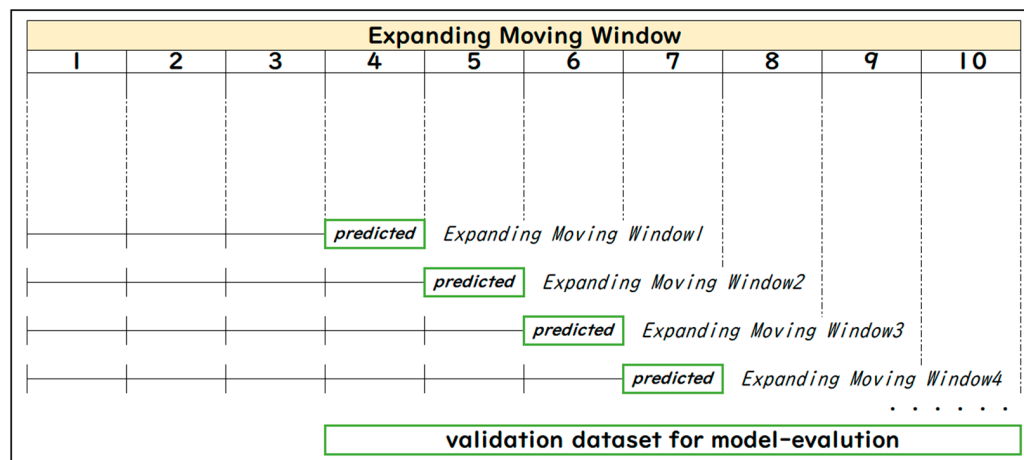


Figure 4. Mechanism of the EMW. Note: The figure illustrates the iterative mechanism of the EMW when adopting an initial window size of three periods for one-period-ahead forecasting.

In terms of the expanding-length window, the initial window size is set to 3363, which is the same as the length of the validation data (there are 3957 observations from 3 February 2003 to 16 December 2020). When iterating the model fitting, the window size increases by one period. For example, the first window spans from 3 February 2003 to 16 July 2017, and is used to estimate 17 July 2017. The framework utilizes the dataset from period 1 to 3363 to train the model, then uses the trained model to forecast period 3364, and incorporates the extended training dataset from period 1 to 3364 to retrain the model. The updated model is then used to predict period 3365. This process is iterated until the last period of the time series. The expanding moving window technique is also employed in the model evaluation as walk-forward testing (Baranochnikov and Ślepaczuk 2022).

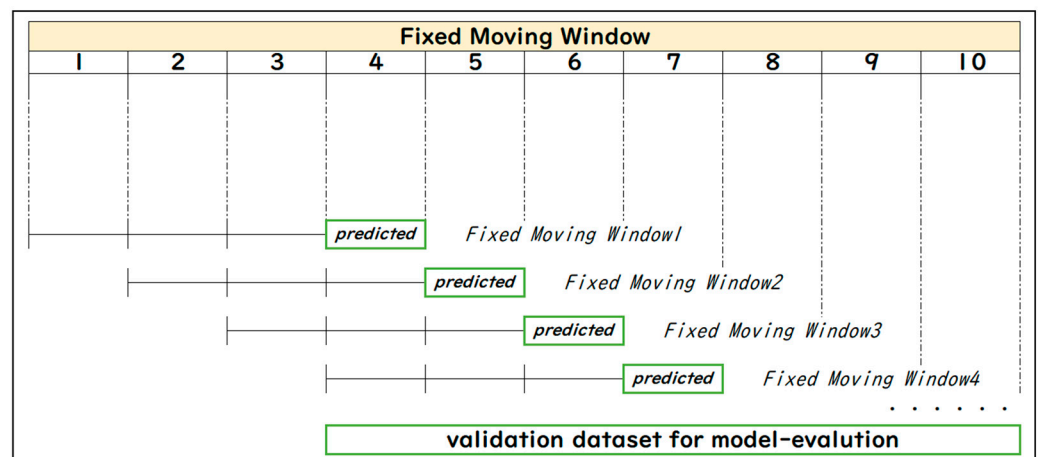


Figure 5. Mechanism of the FMW. Note: The figure illustrates the iterative mechanism of the FMW when adopting an initial window size of three periods for one-period-ahead forecasting.

In terms of the fixed-length window, the window size is determined to be 3363. For instance, the first window spans from 3 February 2003 to 16 July 2017, and is used to estimate 17 July 2017. The model uses the dataset from period 1 to 3363 to train the model and utilizes this trained model to forecast period 3364. Then, the dataset from period 2 to 3364 is used to train the model, and the updated model is used to predict period 3365. This process is iterated until the last period of the time series.

3.2.2. Random Forest (RF)

The RF approach, introduced by Breiman (2001), is an ensemble machine learning method that incorporates multiple decision trees to improve prediction performance. By extending each tree from randomly selected features and building them from the primal sample, the RF method addresses the overfitting problem that can arise when adding more trees to the forest. This approach enhances prediction accuracy.

To maximize the forecasting performance of our model, we conducted a meticulous parameter-tuning process. We optimized several variables to achieve optimal results in our forecasting endeavor. The variables that underwent optimization included *n_estimators* (with values of 100, 200, 300, 400, and 500), *max_depth* (with values of 1, 3, 10, 20, 30, 40, and 50), *bootstrap* (with options of True and False), and *min_samples_leaf* (ranging from 1 to 10). After a thorough evaluation based on error metrics, we selected the following parameter values: *n_estimators* (300), *max_depth* (20), *bootstrap* (True), and *min_samples_leaf* (3). These parameter values were found to yield the best performance in our model, ensuring accurate and reliable forecasting outcomes.

3.2.3. Extreme Gradient Boosting (XGBoost)

XGBoost, an algorithm proposed by Chen and Guestrin (2016), is an ensemble machine learning model that enhances gradient boosting techniques (Friedman 2001). It employs an optimized platform for gradient boosting, leveraging parallel processing, tree pruning, and hardware optimization. XGBoost offers a variety of objective functions, including classification and regression, and combines weaker and simpler learner estimates (such as regression trees) to improve prediction accuracy. The model minimizes a subjective loss function through a penalty term for model complexity (i.e., regression tree functions) and a convex loss function. Iterative learning involves creating new trees and merging them with existing trees.

To enhance the predictive performance of our model, we conducted a meticulous parameter-tuning process. We optimized several variables to achieve optimal results in our forecasting endeavor. The variables that underwent optimization included *n_estimators* (ranging from 100 to 1000 in increments of 100), *max_depth* (with values of 1, 3, 5, and 10),

learning_rate (with values of 0.001 and 0.01), and gamma (with values of 0, 0.001, and 0.01). Subsequently, based on the performance evaluation using error metrics, we selected the following parameter values: n_estimators (1000), max_depth (3), learning_rate (0.01), and gamma (0.01).

3.2.4. Long Short-Term Memory (LSTM)

The LSTM algorithm was first introduced by Hochreiter and Schmidhuber (1997). As a prominent model in deep learning, LSTM exhibits an external loop structure similar to that of RNN and an internal recurrent structure consisting of memory cells. Each memory cell possesses self-connected recurrent weights that interact with three types of gates, ensuring the preservation of signals over multiple time steps without suffering from exploding or vanishing gradients. Similar to RNN, LSTM can utilize more data at each time step, resembling the memory capacity of the LSTM unit. The network utilizes these gates to effectively manage the retention and forgetting of information for subsequent iterations.

To achieve optimal forecasting outcomes, we meticulously tuned the hyperparameters of our model. Various variables underwent optimization, including batch size (ranging from 10 to 200), number of epochs (ranging from 10 to 300), optimization technique (SGD, Adam, RMSprop), learning rate (0.001, 0.01, 0.1), dropout rate (ranging from 0.0 to 0.9), neuron activation function (relu, sigmoid), number of layers (ranging from 1 to 5), and number of neurons (16, 32, 46, 64, 128). During the training of the neural networks, we employed the traditional mean squared error (MSE) loss function, as utilized by Cao et al. (2019), Chimmula and Zhang (2020), and Livieris et al. (2020). This loss function is widely recognized and commonly used in the field. Following a comprehensive evaluation process, we selected the following parameter values that exhibited superior performance: a batch size of 15, 150 epochs, the Adam optimization technique, a learning rate of 0.001, no dropout (dropout rate of 0.0), relu activation function, 3 layers, and 46 neurons. These parameter values were determined to produce the most accurate and reliable forecasting results in our model.

3.2.5. AutoRegressive Integrated Moving Average (ARIMA)

ARIMA was developed in the 1970s by Box and Jenkins (1968) with the aim of mathematically characterizing variations in time series. Non-stationary data need to be differenced until stationarity is achieved, as ARIMA specifically works with stationary data. In ARIMA (p, d, q), where p represents the autoregressive terms, d represents the differencing order, and q represents the lagged errors, the best values for p, d, and q are determined using the Akaike information criterion to fit the data.

In this study, the selection of optimal (p, d, q) values for time series analysis is performed using the auto_arima function in Python. The auto_arima function employs a stepwise search method to minimize the Akaike Information Criteria (AIC). To ensure model parsimony, the maximum values for p and q are set to be less than 5. The determination of the optimal differencing parameter, d, is achieved through the application of the Augmented Dickey-Fuller test.

3.2.6. Wavelet-ARIMA-LSTM (Wavelet-ARIMA-RF/Wavelet-ARIMA-XGB)

The wavelet transform was first introduced by French scientist J. Morlet in 1974 (Morlet et al. 1982). Wavelet decomposition has been widely utilized as a preprocessing approach in various fields such as engineering, time series analysis, and medicine. By applying wavelet decomposition, time series data can be separated into approximation and detail components. In this study, we employ discrete wavelet decomposition (DWD) to decompose the gold futures prices and the euro exchange rate into multiple approximation and detail component series. Unlike previous research, we simplify the analysis by using the decomposed approximation series for forecasting one-period-ahead values using the ARIMA model. We then calculate the residuals and apply the LSTM model to predict the one-period-ahead residuals, and finally combine them.

In summary, the DWD technique is employed to decompose the price time series into linear approximation components and nonlinear residual components. The linear components are predicted using the ARIMA model, while the nonlinear parts are independently forecasted using the LSTM model, taking into account the intrinsic characteristics of these models.

Similarly, in the case of wavelet-ARIMA-RF and wavelet-ARIMA-XGB, the random forest model and extreme gradient boosting are applied, respectively, to predict the nonlinear components.

3.2.7. Seasonal-Decomposition-ARIMA-LSTM

Furthermore, we employed another preprocessing technique, known as traditional seasonal decomposition, for the time series models of gold futures prices and the euro exchange rate. According to the traditional concept of time series decomposition, a series is considered as a composite of level, trend, seasonality, and noise components. In this study, we regard the level, trend, and seasonality components as systematic components since they exhibit consistency or recurrence and can be described and modeled. Conversely, we classify the noise component as non-systematic due to its random variation nature. Diverging from previous research, we utilize the decomposed systematic components, including the trend series and seasonality series, to apply the ARIMA model for forecasting one-period-ahead values. Subsequently, we employ the decomposed non-systematic noise components to apply the LSTM model for predicting one-period-ahead noise, and ultimately aggregate these predicted values.

In summary, the traditional seasonal decomposition method is utilized to decompose the price time series into linear systematic components and nonlinear non-systematic components. The linear components are then forecasted using the ARIMA model, while the nonlinear components are separately predicted using the LSTM model.

3.3. Model Evaluation Measures

3.3.1. Root Mean Squared Error (RMSE)

The discrepancy between the expected and actual values is typically measured using the RMSE. The RMSE is typically computed as follows:

$$RMSE = \sqrt{\frac{\sum_{i=1}^N (x_i - \hat{x}_i)^2}{N}} \quad (2)$$

where N is the number of non-missing data points, x_i is the actual observation time series, and \hat{x}_i is the estimated time series.

3.3.2. Mean Absolute Percentage Error (MAPE)

The accuracy of forecasting models is frequently assessed statistically using the mean absolute percentage error (MAPE). MAPE can be calculated as the average absolute percent error for each time period minus actual values divided by actual values. Generally speaking, the following equation defines MAPE:

$$MAPE = \frac{1}{n} \sum_{i=1}^N \left| \frac{x_i - \hat{x}_i}{x_i} \right| \quad (3)$$

where i = variable, N = number of non-missing data points, x_i = actual observation time series, \hat{x}_i = estimated time series.

This paper defined the MAPE accuracy (%) by $MAPE (\%) = 100 * MAPE$.

3.3.3. Mean Absolute Error (MAE)

The mean absolute error (MAE) is frequently used as a statistical measure of the average magnitude of the errors in a predicted dataset without considering their direction. It is the average over the test sample of the absolute differences between prediction and

actual observation, where all individual differences have equal weight. Generally, MAE is defined by the following equation:

$$MAE = \frac{1}{n} \sum_{i=1}^N |x_i - \hat{x}_i| \tag{4}$$

where i = variable, N = number of non-missing data points, x_i = actual observation time series, \hat{x}_i = estimated time series.

3.3.4. Modified Diebold–Mariano Test

The DM test was originally introduced by Diebold and Mariano (1995). In empirical analyses, when there are two or more time series forecasting models, it is often a challenge to predict which model is more accurate or whether they are equally suitable. This test identifies whether the null hypothesis (i.e., that the competing model holds equivalent forecasting power as the base model) is statistically true. Assuming that the actual values $\{y_t; t = 1, \dots, T\}$, two forecasts $\{\hat{y}_{1t}; t = 1, \dots, T\}$, $\{\hat{y}_{2t}; t = 1, \dots, T\}$, and forecast error ε_{it} are as follows:

$$\varepsilon_{it} = \hat{y}_{it} - y_t, i = 1, 2 \tag{5}$$

where ε_{it} denotes the forecast error and the loss function, $g(\varepsilon_{it})$, which is defined by the following function:

$$g(\varepsilon_{it}) = (\varepsilon_{it})^2 \tag{6}$$

Then, the loss differential d_t is expressed as follows:

$$d_t = g(\varepsilon_{1t}) - g(\varepsilon_{2t}) \tag{7}$$

Correspondingly, the statistic for the DM test is expressed using the following formula:

$$DM = \frac{\bar{d}}{\sqrt{\frac{s}{N}}} \tag{8}$$

where \bar{d} , s , and N denote the mean loss differential, the variation of d_t , and the number of data points, respectively.

The null hypothesis is $H_0 : E[d_t] = 0, \forall t$, meaning that the two forecast models hold equivalent forecasting performance. Meanwhile, the alternative hypothesis is $H_1 : E[d_t] \neq 0, \forall t$, which represents the difference in accuracy between these two forecasts. Under the null hypothesis, the statistics for the DM test are asymptotically $N(0, 1)$ normally distributed. The null hypothesis would be rejected when $DM > 1.96$.

Harvey et al. (1997) proposed a modified DM test. They suggested that the modified DM test is more suitable when using a small sample. The statistic for the modified DM test is expressed as follows:

$$DM^* = \sqrt{[n + 1 - 2h + h(h - 1)]n^{-1}}DM \tag{9}$$

where h represents the horizon and DM refers to the original DM statistic. Here, we predicted one-period-ahead; hence, $h = 1$; hence,

$$DM^* = \sqrt{(n - 1)n^{-1}}DM \tag{10}$$

Concerning how to interpret the DM test statistic results, since we set $g(\varepsilon_{1t})$ as the target model, $g(\varepsilon_{2t})$ as the base model, the numerator is (target-base), therefore, if the DM test statistic is negative, that means the target model has a smaller variance than the base model; hence, the prediction performance of the target model is better than the base model. The p -value denotes the significance of this statistic.

4. Results

4.1. Empirical Results

4.1.1. Prediction Results of SI Dataset and Large Dataset

Firstly, this subsection presents the prediction performance results of the sentiment dataset and the large dataset to verify whether the sentiment dataset could replace the large dataset when predicting commodity gold prices and the euro foreign exchange rate.

Table 2 displays the prediction outcomes for gold futures prices utilizing the sentiment indicator dataset, while Table 3 presents the prediction results for gold futures prices employing the large dataset. Likewise, Table 4 lists the prediction results for the euro foreign exchange rate based on the sentiment indicator dataset, and Table 5 showcases the prediction results for the euro foreign exchange rate utilizing the large dataset. Overall, the prediction results indicate that the sentiment indicator dataset generally exhibits better forecasting performance than the large dataset. When comparing the performance metrics, namely RMSE, MAPE, and MSE, between the two datasets, it becomes evident that the fixed moving window LSTM approach using the SI dataset outperforms the alternative dataset and models considered. This finding suggests that combining the sentiment indicator with the moving window LSTM machine learning model yields the best results for predicting gold futures prices and euro exchange rates. These results align with the outcomes of previous studies by Plakandaras et al. (2015), Nwosu et al. (2021), and Dunis and Williams (2002), which suggest that neural network models or their proposed approaches, particularly when combined with neural networks, offer more accurate forecasts compared to other models. Furthermore, these results provide additional evidence supporting the superiority of the LSTM model’s complex loop structure. Turning to the forecasting results using the large dataset, the moving window RF results demonstrate the best performance. This may be attributed to the use of a large indicator dataset, which allows the RF classifier to effectively enhance the predictive power. Although our study employs a different data source for sentiment analysis compared to previous research (Naeem et al. 2021), our empirical results broadly align with the findings of Li et al. (2016) and Naeem et al. (2021) in terms of predicting gold futures and euro exchange rates, thus indicating that the sentiment dataset can serve as a viable substitute for the large dataset.

Table 2. Results of the SI dataset for gold futures prices.

Dataset	Evaluation	RF_EMW	RF_FMW	XGBoost_EMW	XGBoost_FMW	LSTM_EMW	LSTM_FMW
SI dataset	RMSE	10.3122	10.4261	9.9711	9.9852	11.1461	9.3283
	MAPE	0.5810	0.5832	0.5480	0.5480	0.6160	0.5130
	MSE	7.7159	7.7159	7.3015	7.3056	8.2074	6.8072

Note: RF represents random forest. XGBoost denotes eXtreme gradient boosting. LSTM denotes long short-term memory. The underline followed by EMW denotes the expanding moving window technique, while the underline followed by FMW denotes the fixed moving window technique. RMSE denotes the root mean squared error. MAPE denotes the mean absolute percentage error. MAE denotes the mean absolute error. The best performance in this set of prediction results is shown in bold.

Table 3. Results of the large dataset for gold futures prices.

Dataset	Evaluation	RF_EMW	RF_FMW	XGBoost_EMW	XGBoost_FMW	LSTM_EMW	LSTM_FMW
Large dataset	RMSE	9.8462	9.8752	10.6259	10.6373	14.0267	11.1016
	MAPE	0.5450	0.5470	0.5910	0.5930	0.8230	0.6340
	MSE	7.2544	7.2544	7.9100	7.9534	10.8288	8.3960

Note: RF represents random forest. XGBoost denotes eXtreme gradient boosting. LSTM denotes long short-term memory. The underline followed by EMW denotes the expanding moving window technique, while the underline followed by FMW denotes the fixed moving window technique. RMSE denotes the root mean squared error. MAPE denotes the mean absolute percentage error. MAE denotes the mean absolute error. The best performance in this set of prediction results is shown in bold.

Table 4. Results of the SI dataset for euro foreign exchange rate.

Dataset	Evaluation	RF_EMW	RF_FMW	XGBoost_EMW	XGBoost_FMW	LSTM_EMW	LSTM_FMW
SI dataset	RMSE	0.550	0.553	0.596	0.603	0.479	0.47384
	MAPE	0.370	0.372	0.396	0.400	0.322	0.3180
	MSE	0.430	0.430	0.460	0.465	0.374	0.3693

Note: RF represents random forest. XGBoost denotes eXtreme gradient boosting. LSTM denotes long short-term memory. The underline followed by EMW denotes the expanding moving window technique, while the underline followed by FMW denotes the fixed moving window technique. RMSE denotes the root mean squared error. MAPE denotes the mean absolute percentage error. MAE denotes the mean absolute error. The best performance in this set of prediction results is shown in bold.

Table 5. Results of the large dataset for euro foreign exchange rate.

Dataset	Evaluation	RF_EMW	RF_FMW	XGBoost_EMW	XGBoost_FMW	LSTM_EMW	LSTM_FMW
Large dataset	RMSE	0.518	0.521	0.775	0.870	0.677	0.583
	MAPE	0.343	0.346	0.484	0.546	0.412	0.395
	MSE	0.399	0.399	0.563	0.636	0.476	0.458

Note: RF represents random forest. XGBoost denotes eXtreme gradient boosting. LSTM denotes long short-term memory. The underline followed by EMW denotes the expanding moving window technique, while the underline followed by FMW denotes the fixed moving window technique. RMSE denotes the root mean squared error. MAPE denotes the mean absolute percentage error. MAE denotes the mean absolute error. The best performance in this set of prediction results is shown in bold.

4.1.2. Prediction Results of ARIMA Model

However, when comparing with the classical statistical model, ARIMA, whether the conclusion holds robustness needs to be investigated. Therefore, we conducted the simple prediction by ARIMA, and the lags were chosen using the Akaike Information Criteria (AIC). The forecasting results are presented in Table 6.

Table 6. Results of the ARIMA for gold futures prices and the euro foreign exchange rate.

Evaluation	Gold	Euro
RMSE	9.2658	0.47388
MAPE	0.5090	0.3170
MSE	6.7591	0.3687

Note: RMSE denotes the root mean squared error. MAPE denotes the mean absolute percentage error. MAE denotes the mean absolute error.

Based on the above results, we are pleasantly surprised by the effectiveness of the powerful yet simple statistical model, ARIMA, in predicting time series. This finding aligns with the research reported by He (2018). However, it contradicts the studies conducted by Siami-Namini et al. (2018) and Siami-Namini et al. (2019). These results suggest that simplicity may be the key when it comes to designing prediction models for time series, despite the prevalence of complex models and fancy datasets. In contrast to the findings of Nwosu et al. (2021) and Dunis and Williams (2002), our results indicate that it is worth considering the use of simple traditional models in the design of prediction models.

4.1.3. Prediction Results of Proposed Approaches

However, it is worth noting that machine learning and deep learning models have been extensively validated in numerous studies for their superior effectiveness and accuracy in predicting time series compared to ARIMA models. Therefore, it is necessary to further verify the robustness of the simple statistical model, ARIMA. Inspired by Abdulrahman et al. (2021) and others, we propose a triple combination of wavelet-ARIMA-LSTM, wavelet-ARIMA-RF, and wavelet-ARIMA-XGB models, as well as the seasonal-decomposition-ARIMA-LSTM approach, to investigate this objective. The prediction results are summarized in Tables 7 and 8.

Table 7. Results of the proposed approaches for gold futures prices.

Evaluation	SeasonalDecomposition_ ARIMA_LSTM	Wavelet_ ARIMA_LSTM	Wavelet_ ARIMA_XGB	Wavelet_ ARIMA_RF	ARIMA	LSTM	XGB	RF
RMSE	3.3916	8.4439	5.4610	5.4610	9.2658	12.2605	10.0311	10.8282
MAPE	0.0020	0.4840	0.3060	0.3060	0.5090	0.7670	0.5420	0.6140
MSE	2.6869	6.4376	4.0516	4.0516	6.7591	9.9841	7.2283	8.1704

Note: RF represents random forest. XGBoost denotes eXtreme gradient boosting. LSTM denotes long short-term memory. RMSE denotes the root mean squared error. MAPE denotes the mean absolute percentage error. MAE denotes the mean absolute error. The best performance in this set of prediction results is shown in bold. SeasonalDecomposition denotes the seasonal decomposition. Wavelet represents the wavelet decomposition.

Table 8. Results of the proposed approaches for euro foreign exchange rate.

Evaluation	SeasonalDecomposition_ ARIMA_LSTM	Wavelet_ ARIMA_LSTM	Wavelet_ ARIMA_XGB	Wavelet_ ARIMA_RF	ARIMA	LSTM	XGB	RF
RMSE	0.1632	0.4083	0.1813	0.3443	0.4739	0.5578	0.5952	0.6526
MAPE	0.1120	0.2687	0.1200	0.2400	0.3170	0.3960	0.3950	0.4410
MSE	0.1298	0.3122	0.1389	0.2784	0.3687	0.4573	0.4595	0.5122

Note: RF represents random forest. XGBoost denotes eXtreme gradient boosting. LSTM denotes long short-term memory. RMSE denotes the root mean squared error. MAPE denotes the mean absolute percentage error. MAE denotes the mean absolute error. The best performance in this set of prediction results is shown in bold. SeasonalDecomposition denotes the seasonal decomposition. Wavelet represents the wavelet decomposition.

Based on the results presented in Tables 7 and 8, Figures 6 and 7, our proposed triple-combination approach demonstrates superior prediction accuracy compared to individual ARIMA, machine learning, and deep learning approaches. This suggests that by decomposing time series into linear and nonlinear components and combining classical statistical models with machine learning approaches, we achieve more precise predictions. However, the best performing approach for both object time series, namely gold futures prices and the euro foreign exchange rate, is the SeasonalDecomposition_ARIMA_LSTM model. It is followed by Wavelet_ARIMA_XGB and Wavelet_ARIMA_RF. This finding suggests that the systematic and non-systematic decomposition combined with the ARIMA and LSTM models for predicting commodity prices and foreign exchange rates is preferable. These results align with previous studies (Chang et al. 2019; Chen and Wang 2019; Liu et al. 2018; Ma et al. 2019; Moustafa and Khodairy 2023), further supporting the effectiveness of the integrated multiple-model approach in prediction. Our empirical forecasting results provide additional evidence that the multiple-model integrated approach performs better in prediction.



GOLD Future Prices and SeasonalDecomposition_ARIMA_LSTM Predicted Prices

Figure 6. Prediction result plotting for gold future prices. Note: This figure illustrates the gold futures prices and the predicted values.

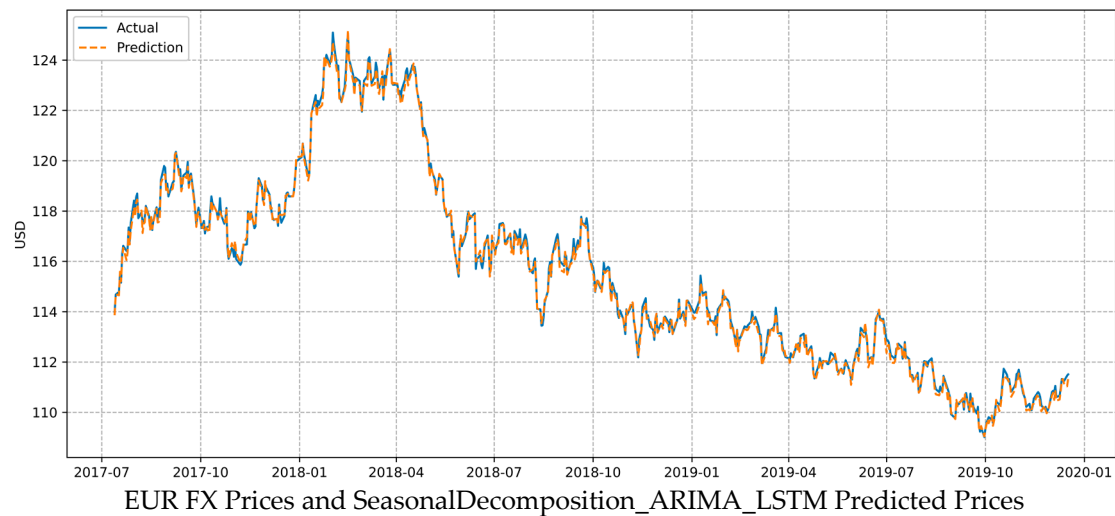


Figure 7. Prediction result plotting for euro foreign exchange rates. Note: This figure illustrates the euro foreign exchange rates and the predicted values.

In summary, first, the combination of the sentiment indicator with the fixed moving window LSTM machine learning model produces the best prediction results compared to the large dataset. This result demonstrates that sentiment indicators obtained through sentiment analysis outperform the large dataset in terms of prediction ability and can be utilized as a better alternative independent predictor. Second, based on the prediction results, the traditional and classical ARIMA model surprisingly outperforms both the sentiment indicator dataset and the large dataset combined with machine learning techniques. Finally, our proposed triple-combination techniques are superior to both machine learning models and the traditional statistical ARIMA model in terms of commodity price and foreign exchange rate prediction performance. The top three performing forecasting methods are the seasonal-decomposition_ARIMA_LSTM, the wavelet_ARIMA_XGB, and the wavelet_ARIMA_RF. In the first step, these approaches decompose the data into linear and nonlinear components by adopting seasonal decomposition or wavelet transformation. In the second step, they use the ARIMA model to predict the linear part and machine learning or deep learning models to predict the nonlinear part.

4.2. Model Evaluation Results

4.2.1. Walk-Forward Testing Results

In this study, we employ the walk-forward testing method as the chosen back-testing technique to validate the effectiveness of the proposed triple-combination approaches. To evaluate the performance of these models, we adopt an expanding moving window approach, focusing on the last 50 observations. The testing procedure involves conducting separate walk-forward tests on each decomposition component, followed by aggregating the results and comparing the error metrics against those obtained from the ARIMA model.

As we present in Tables 9 and 10, the walk-forward testing results for gold futures prices and euro foreign exchange provide robust estimations for evaluating the effectiveness of our proposed triple-combination approaches. These results offer valuable insights into the performance and reliability of the models in predicting the respective market dynamics.

Table 9. Results of Walk-Forward Testing for gold futures prices.

Evaluation	SeasonalDecomposition_ARIMA_LSTM	Wavelet_ARIMA_LSTM	Wavelet_ARIMA_XGB	Wavelet_ARIMA_RF	ARIMA
RMSE	2.8765	3.1308	3.5565	5.0426	9.2658
MAPE	0.1445	0.1638	0.1868	0.2642	0.5090
MSE	2.1304	2.4265	2.7682	3.9170	6.7591

Note: RF represents random forest. XGBoost denotes eXtreme gradient boosting. LSTM denotes long short-term memory. SeasonalDecomposition denotes the seasonal decomposition. Wavelet represents the wavelet decomposition. RMSE denotes the root mean squared error. MAPE denotes the mean absolute percentage error. MAE denotes the mean absolute error.

Table 10. Results of Walk-Forward Testing for euro foreign exchange rate.

Evaluation	SeasonalDecomposition_ARIMA_LSTM	Wavelet_ARIMA_LSTM	Wavelet_ARIMA_XGB	Wavelet_ARIMA_RF	ARIMA
RMSE	0.1263	0.1028	0.1206	0.3142	0.4739
MAPE	0.0937	0.0762	0.0886	0.2226	0.3172
MSE	0.1037	0.0844	0.0980	0.2466	0.3687

Note: RF represents random forest. XGBoost denotes eXtreme gradient boosting. LSTM denotes long short-term memory. SeasonalDecomposition denotes the seasonal decomposition. Wavelet represents the wavelet decomposition. RMSE denotes the root mean squared error. MAPE denotes the mean absolute percentage error. MAE denotes the mean absolute error.

4.2.2. Diebold–Mariano Test Results

The Diebold–Mariano test is conducted to assess the predictive superiority of the triple-combination approaches compared to the ARIMA models. We present the results of this test in Tables 10 and 11, offering insights into the relative performance of the proposed approaches. The DM test results for both gold futures prices and euro foreign exchange rates are analyzed.

Table 11. DM test results of gold futures prices.

Target Approach	Base Model (ARIMA)	
	DM Test	p-Value
SeasonalDecomposition_ARIMA_LSTM	−9.9779	0.000
Wavelet_ARIMA_LSTM	−9.4216	0.000
Wavelet_ARIMA_XGB	−9.9468	0.000
Wavelet_ARIMA_RF	−7.1182	0.000

Note: DM test indicates the modified Diebold–Mariano test statistic. RF represents random forest. XGBoost denotes eXtreme gradient boosting. LSTM denotes long short-term memory. SeasonalDecomposition denotes the seasonal decomposition. Wavelet represents the wavelet decomposition. RMSE denotes the root mean squared error. MAPE denotes the mean absolute percentage error. MAE denotes the mean absolute error.

From the results presented in Tables 11 and 12, it is noteworthy that the proposed triple-combination approaches demonstrate a significant outperformance over the classical statistical model, the ARIMA model.

Table 12. DM test results of euro foreign exchange rate.

Target Approach	Base Model (ARIMA)	
	DM Test	p-Value
SeasonalDecomposition_ARIMA_LSTM	−12.9469	0.000
Wavelet_ARIMA_LSTM	−4.5330	0.000
Wavelet_ARIMA_XGB	−12.6462	0.000
Wavelet_ARIMA_RF	−6.3385	0.000

Note: DM test indicates the modified Diebold–Mariano test statistic. RF represents random forest. XGBoost denotes eXtreme gradient boosting. LSTM denotes long short-term memory. SeasonalDecomposition denotes the seasonal decomposition. Wavelet represents the wavelet decomposition. RMSE denotes the root mean squared error. MAPE denotes the mean absolute percentage error. MAE denotes the mean absolute error.

5. Conclusions and Policy Implications

As highlighted by Naeem et al. (2021) and Li et al. (2016), the rapid advancement of the Internet and big data technology has led to an abundance of online data, including textual data from sources such as Twitter and news releases, which can help to identify influential factors in specific markets. Motivated by this, our study aims to examine whether the sentiment indicator dataset obtained through sentiment analysis of unstructured online news headlines can serve as a substitute for the large dataset comprising various indicators in predicting commodity prices and foreign exchange rates.

In our empirical analysis, we employ sentiment analysis using the Python natural language processing library to process news headlines from ABC, which consists of 1,226,258 news headlines, to derive a sentiment indicator. Additionally, we collect 30 additional indicators to construct the large dataset. Subsequently, we utilize this sentiment indicator in conjunction with moving window machine learning and deep learning models, namely RF, XGBoost, and LSTM, to forecast commodity gold futures prices and the euro exchange rate. Alongside comparing the prediction performance of the datasets, we also conduct a prediction comparison between the classical statistical model, ARIMA, and time-varying parameter machine learning models.

Based on the results of the model comparisons, we cannot conclude that sentiment indicators combined with machine learning outperform the ARIMA model. However, from an alternative perspective, we propose triple-combination approaches that involve decomposing the time series data into linear and nonlinear components and subsequently forecasting the linear component using the robust statistical model, ARIMA, and the nonlinear component using machine learning models such as LSTM, XGB, and RF. This research sheds light on the issue of comparing the out-of-sample superiority of our proposed triple-combination approaches for foreign exchange rate prediction with the traditional powerful statistical model, ARIMA. Furthermore, we conduct walk-forward testing to validate the triple-combination approaches and employ the modified Diebold–Mariano test statistic to investigate statistically significant differences between the proposed approach and the ARIMA model.

The study's primary conclusions are as follows: Firstly, the combination of the sentiment indicator with the moving window LSTM machine learning model demonstrates the best forecasting performance. These findings align with previous studies conducted by Plakandaras et al. (2015), Nwosu et al. (2021), and Dunis and Williams (2002). Secondly, the sentiment indicator dataset used by deep learning and moving window machine learning models does not surpass the classical ARIMA model, consistent with the findings reported by He (2018). This result contradicts the studies conducted by Siami-Namini et al. (2018) and Siami-Namini et al. (2019). Thirdly, the proposed triple-combination methods, which expand upon and derive from the approaches of Chang et al. (2019), Chen and Wang (2019), Liu et al. (2018), Ma et al. (2019), and Moustafa and Khodairy (2023), exhibit superior performance in predicting commodity prices and foreign exchange rates compared to both machine learning models and the ARIMA model. The seasonal-decomposition ARIMA-LSTM, wavelet-ARIMA-XGB, and wavelet-ARIMA-RF demonstrate the top three forecasting performances based on error metrics, walk-forward testing results, and Diebold–Mariano test results. In the first step, the data are decomposed into linear and non-linear components using wavelet transformation or seasonal decomposition. In the second step, the linear component is predicted using ARIMA, while the non-linear component is predicted using machine learning or deep learning models. Lastly, in addition to the aforementioned findings, the comparison of results between the sentiment indicator dataset and the large dataset indicate that sentiment indicators obtained through sentiment analysis possess superior forecasting capabilities compared to the large dataset consisting of various indicators. Consequently, they can be utilized as better alternative predictors. Our empirical results generally align with the findings of Li et al. (2016) and Naeem et al. (2021) in terms of predicting gold futures and euro exchange rates, further highlighting the potential of the sentiment dataset to enhance forecasting in time series prediction.

To the best of our knowledge, this study presents a pioneering investigation into the potential of sentiment indicators as a substitute for extensive datasets in forecasting commodity prices and foreign exchange rates. The novelty lies in proposing a novel integration of machine learning models, statistical models, and data decomposition techniques to enhance price predictions in these markets. Importantly, the results validate the superior accuracy of the proposed triple-combination approach compared to individual models. Furthermore, these findings offer valuable insights for investors and policymakers, providing them with fresh perspectives, predictive tools, and alternative forecasting approaches.

For investors, the research offers fresh perspectives on forecasting commodity prices and foreign exchange rates. It introduces new predictive tools and alternative approaches that enhance their decision-making processes and potentially lead to more accurate forecasts. Additionally, precise prediction of gold prices and euro exchange rates is crucial for informing hedging strategies aimed at mitigating risks arising from currency fluctuations.

For policymakers, these findings play a vital role in making informed investment decisions. Gold is widely utilized as a means to hedge against inflation and market volatility, and fluctuations in the euro exchange rate have a substantial impact on the costs and risks associated with international transactions as the second most traded currency globally. Moreover, improving the accuracy of gold price predictions is crucial for central banks that maintain gold reserves as a safeguard against currency fluctuations and as a store of value. Given that the euro is a major reserve currency used in international transactions and investments, precise prediction of euro exchange rates can bolster financial stability. Furthermore, gold prices and euro exchange rates are closely intertwined with the international economy and play a pivotal role in informing government policy decisions regarding trade, monetary policy, and capital flows. Therefore, our findings contribute to economic forecasting, empowering policymakers and investors to leverage these predictions for informed decision making, ensuring they are well-prepared to navigate and respond to evolving economic conditions.

Despite these findings, this study has its limitations. Since we only employ RF, XGB, and LSTM methods to compare forecasts with the ARIMA model, we cannot conclusively determine that ARIMA is superior to other machine learning and deep learning models. Further verification is necessary to address this point. Additionally, there are numerous other data decomposition methods that require testing to validate the conclusions.

In future research, it is recommended to explore alternative data decomposition methods, as well as additional machine learning and deep learning techniques, to expand the investigation to major commodity prices and currency exchange rates. This will help validate the rationality and robustness of the proposed approaches' superiority. Furthermore, considering the potential of the sentiment indicator as a promising alternative dataset, empirical testing is planned to assess whether incorporating the proposed approaches with the additional sentiment indicator can further enhance the forecasting accuracy for commodity prices and foreign exchange rates.

Author Contributions: Investigation, J.S.; writing—original draft preparation, J.S.; writing—review and editing, S.H.; project administration, S.H.; funding acquisition, S.H. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by JSPS KAKENHI [(grant number: 22K01424)].

Data Availability Statement: Not applicable.

Acknowledgments: We are grateful to four anonymous reviewers for their helpful comments and suggestions.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

The descriptions of and sources of the data are presented in Table A1.

Table A1. Descriptions and sources of the indicators used in this study.

Variable	Description	Source
EUR	Euro against the US dollar	Investing.com
CAD	Canadian dollar against the US dollar	Investing.com
JPY	Japanese yen against the US dollar	Investing.com
WTIf	WTI Crude Oil futures prices	Bloomberg
Brent_oil	Brent Crude Oil futures prices	Investing.com
Henryhub_gas	Henry Hub Natural Gas futures prices	Bloomberg
SP500	Standard & Poor's 500 Stock Index	FRB ¹
FTSE100	The Financial Times Stock Exchange Group:London Stock Exchange	FRB ¹
NASDAQ	NASDAQ Composite Index	FRB ¹
HangSeng	Hong Kong Hang Seng Composite stock market index	Macrotrends
CAC40	France's CAC 40 stock market index	Macrotrends
GSPTSE	Canadian S&P/TSX Composite Index	Investing.com
US10_Bond	US 10-Year Treasury Constant Maturity Rate	Yahoo! Finance
UK10_Bond	United Kingdom 10-Year Bond Yield	Investing.com
Germany10_Bond	Germany 10-Year Bond Yield	Investing.com
DAX	Germany's DAX 30 stock market index	Macrotrends
NIKKEI	Tokyo Stock Exchange:Nikkei index	FRB ¹
Gold	Gold futures prices	Bloomberg
TWUSDI	Trade Weighted U.S. Dollar Index	FRB ¹
FederalFunds	Federal Funds Rate	Macrotrends
CORN	Corn futures prices	Datastream ²
WHEAT	Wheat futures prices	Datastream ²
RSI	Relative Strength Index	Calculated
ma7	7-days Moving Average	Calculated
ma21	21-days Moving Average	Calculated
26ema	26-days Exponential Weighted Moving Average	Calculated
12ema	12-days Exponential Weighted Moving Average	Calculated
MACD	Moving Average Convergence/Divergence oscillator	Calculated
20sd	20-days Standard Deviation	Calculated
upper_band	Bollinger Bands	Calculated
lower_band	Bollinger Bands	Calculated
ema	Exponential Moving Average	Calculated

Note: ¹ Federal Reserve Bank. ² Thomson Reuters Datastream.

References

- Abdulrahman, Umar Farouk Ibn, Najim Ussiph, and Benjamin Hayfron-Acquah. 2021. A Hybrid Arima-Lstm Model for Stock Price Prediction. *International Journal of Computer Engineering and Information Technology* 12: 48–51. Available online: <https://www.proquest.com/openview/288bcbf49b187672d89c8a93865cc9d0/1?pq-origsite=gscholar&cbl=2044551> (accessed on 25 April 2023).
- Aizenman, Joshua, and Kenta Inoue. 2013. Central Banks and Gold Puzzles. *Journal of the Japanese and International Economies* 28: 69–90. [CrossRef]
- Amat, Christophe, Tomasz Michalski, and Gilles Stoltz. 2018. Fundamentals and Exchange Rate Forecastability with Simple Machine Learning Methods. *Journal of International Money and Finance* 88: 1–24. [CrossRef]

- Amato, Jeffery D., Andrew J. Filardo, Gabriele Galati, Goetz von Peter, and Feng Zhu. 2005. Research on Exchange Rates and Monetary Policy: An Overview. *SSRN Electronic Journal*. [CrossRef]
- Bakay, Melahat Sevgül, and Ümit Ağbulut. 2021. Electricity Production Based Forecasting of Greenhouse Gas Emissions in Turkey with Deep Learning, Support Vector Machine and Artificial Neural Network Algorithms. *Journal of Cleaner Production* 285: 125324. [CrossRef]
- Baranochnikov, Illia, and Robert Ślepaczuk. 2022. *A Comparison of LSTM and GRU Architectures with the Novel Walk-Forward Approach to Algorithmic Investment Strategy*. No. 2022-21. Warsaw: QFRG.
- Bedi, Punam, and Purnima Khurana. 2019. Sentiment Analysis Using Fuzzy-Deep Learning. In *Proceedings of ICETIT 2019: Emerging Trends in Information Technology*. Berlin: Springer, pp. 246–57. [CrossRef]
- Blose, Laurence E. 2010. Gold Prices, Cost of Carry, and Expected Inflation. *Journal of Economics and Business* 62: 35–47. [CrossRef]
- Bollen, Johan, Huina Mao, and Xiaojun Zeng. 2011. Twitter Mood Predicts the Stock Market. *Journal of Computational Science* 2: 1–8. [CrossRef]
- Bouktif, Salah, Ali Fiaz, Ali Ouni, and Mohamed Serhani. 2018. Optimal Deep Learning LSTM Model for Electric Load Forecasting Using Feature Selection and Genetic Algorithm: Comparison with Machine Learning Approaches. *Energies* 11: 1636. [CrossRef]
- Box, George E. P., and Gwilym M. Jenkins. 1968. Some Recent Advances in Forecasting and Control. *Applied Statistics* 17: 91. [CrossRef]
- Breiman, Leo. 2001. Random forests. *Machine Learning* 45: 5–32. [CrossRef]
- Cao, Jian, Zhi Li, and Jian Li. 2019. Financial Time Series Forecasting Model Based on CEEMDAN and LSTM. *Physica A: Statistical Mechanics and Its Applications* 519: 127–39. [CrossRef]
- Chang, Zihan, Yang Zhang, and Wenbo Chen. 2019. Electricity Price Prediction Based on Hybrid Model of Adam Optimized LSTM Neural Network and Wavelet Transform. *Energy* 187: 115804. [CrossRef]
- Chatzis, Sotirios P., Vassilis Siakoulis, Anastasios Petropoulos, Evangelos Stavroulakis, and Nikos Vlachogiannakis. 2018. Forecasting Stock Market Crisis Events Using Deep and Statistical Machine Learning Techniques. *Expert Systems with Applications* 112: 353–71. [CrossRef]
- Chen, Tianqi, and Carlos Guestrin. 2016. Xgboost: A scalable tree boosting system. Paper presented at the 22nd ACM Sigkdd International Conference on Knowledge Discovery and Data Mining, Washington DC, USA, August 14–18; pp. 785–94.
- Chen, Yuwei, and Kaizhi Wang. 2019. Prediction of Satellite Time Series Data Based on Long Short Term Memory-Autoregressive Integrated Moving Average Model (LSTM-ARIMA). Paper presented at THE 2019 IEEE 4th International Conference on Signal and Image Processing (ICSIP), Wuxi, China, July 19–21.
- Chimmula, Vinay Kumar Reddy, and Lei Zhang. 2020. Time Series Forecasting of COVID-19 Transmission in Canada Using LSTM Networks. *Chaos, Solitons & Fractals* 135: 109864. [CrossRef]
- Chua, Jess, and Richard S. Woodward. 1982. Gold as an inflation hedge: A comparative study of six major industrial countries. *Journal of Business Finance & Accounting* 9: 191–97. [CrossRef]
- Darley, Olufunke G., Abayomi I. O. Yussuff, and Adetokunbo A. Adenowo. 2021. Price Analysis and Forecasting for Bitcoin Using Auto Regressive Integrated Moving Average Model. *Annals of Science and Technology* 6: 47–56. [CrossRef]
- Das, Sushree, Ranjan Kumar Behera, Mukesh Kumar, and Santanu Kumar Rath. 2018. Real-Time Sentiment Analysis of Twitter Streaming Data for Stock Prediction. *Procedia Computer Science* 132: 956–64. [CrossRef]
- Dave, Emmanuel, Albert Leonardo, Marethia Jeanice, and Novita Hanafiah. 2021. Forecasting Indonesia Exports Using a Hybrid Model ARIMA-LSTM. *Procedia Computer Science* 179: 480–87. [CrossRef]
- Deeney, Peter, Mark Cummins, Michael Dowling, and Adam Birmingham. 2015. Sentiment in Oil Markets. *International Review of Financial Analysis* 39: 179–85. [CrossRef]
- Diebold, Francis X, and Robert S Mariano. 1995. Comparing Predictive Accuracy. *Journal of Business and Economic Statistics* 13: 253–63. [CrossRef]
- Dunis, Christian, and Mark Williams. 2002. Modelling and trading the EUR/USD exchange rate: Do neural network models perform better? *Derivatives Use, Trading and Regulation* 8: 211–39.
- Farsi, Behnam, Manar Amayri, Nizar Bouguila, and Ursula Eicker. 2021. On Short-Term Load Forecasting Using Machine Learning Techniques and a Novel Parallel Deep LSTM-CNN Approach. *IEEE Access* 9: 31191–212. [CrossRef]
- Friedman, Jerome H. 2001. Greedy function approximation: A gradient boosting machine. *Annals of Statistics* 29: 1189–232. [CrossRef]
- Guo, Chenkai, Xiaoyu Yan, and Yan Li. 2020. Prediction of Student Attitude towards Blended Learning Based on Sentiment Analysis. Paper presented at THE 2020 9th International Conference on Educational and Information Technology, Oxford, UK, February 11–13; pp. 228–33.
- Harvey, David, Stephen Leybourne, and Paul Newbold. 1997. Testing the Equality of Prediction Mean Squared Errors. *International Journal of Forecasting* 13: 281–91. [CrossRef]
- He, Xin James. 2018. Crude Oil Prices Forecasting: Time Series vs. SVR Models. *Journal of International Technology and Information Management* 27: 25–42. [CrossRef]
- Henry, Elaine, and Andrew J. Leone. 2016. Measuring Qualitative Information in Capital Markets Research: Comparison of Alternative Methodologies to Measure Disclosure Tone. *The Accounting Review* 91: 153–78. [CrossRef]
- Hochreiter, Sepp, and Jürgen Schmidhuber. 1997. Long Short-Term Memory. *Neural Computation* 9: 1735–80. [CrossRef]
- Ito, Tomoki, Kota Tsubouchi, Hiroki Sakaji, Kiyoshi Izumi, and Tatsuo Yamashita. 2019. CSNN: Contextual Sentiment Neural Network. Paper presented at International Conference on Data Mining, Beijing, China, November 8–11.

- Ito, Tomoki, Kota Tsubouchi, Hiroki Sakaji, Tatsuo Yamashita, and Kiyoshi Izumi. 2020. Contextual Sentiment Neural Network for Document Sentiment Analysis. *Data Science and Engineering* 5: 180–92. [CrossRef]
- Júnior, Domingos S. de O. Santos, João F. L. de Oliveira, and Paulo S. G. de Mattos Neto. 2019. An Intelligent Hybridization of ARIMA with Machine Learning Models for Time Series Forecasting. *Knowledge-Based Systems* 175: 72–86. [CrossRef]
- Kulkarni, Rohit. 2018. *A Million News Headlines*. V6. Cambridge: Harvard Dataverse. Available online: <https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/SYBGZL> (accessed on 25 April 2023).
- Latief, Rashid, and Lin Lefen. 2018. The Effect of Exchange Rate Volatility on International Trade and Foreign Direct Investment (FDI) in Developing Countries along ‘One Belt and One Road’. *International Journal of Financial Studies* 6: 86. [CrossRef]
- Li, Hongmin, Jianzhou Wang, and Hufang Yang. 2020. A Novel Dynamic Ensemble Air Quality Index Forecasting System. *Atmospheric Pollution Research* 11: 1258–70. [CrossRef]
- Li, Jian, Zhenjing Xu, Lean Yu, and Ling Tang. 2016. Forecasting Oil Price Trends with Sentiment of Online News Articles. *Procedia Computer Science* 91: 1081–87. [CrossRef]
- Liu, Hui, Xi-wei Mi, and Yan-fei Li. 2018. Wind Speed Forecasting Method Based on Deep Learning Strategy Using Empirical Wavelet Transform, Long Short Term Memory Neural Network and Elman Neural Network. *Energy Conversion and Management* 156: 498–514. [CrossRef]
- Liu, Xiaolei, Zi Lin, and Ziming Feng. 2021. Short-Term Offshore Wind Speed Forecast by Seasonal ARIMA—A Comparison against GRU and LSTM. *Energy* 227: 120492. [CrossRef]
- Livieris, Ioannis E., Emmanuel Pintelas, and Panagiotis Pintelas. 2020. A CNN–LSTM Model for Gold Price Time-Series Forecasting. *Neural Computing and Applications* 32: 17351–60. [CrossRef]
- Luo, Zhaojie, Xiaojing Cai, Katsuyuki Tanaka, Tetsuya Takiguchi, Takuji Kinkyo, and Shigeyuki Hamori. 2019. Can We Forecast Daily Oil Futures Prices? Experimental Evidence from Convolutional Neural Networks. *Journal of Risk and Financial Management* 12: 9. [CrossRef]
- Ma, Rui, Zhongliang Li, Elena Breaz, Chen Liu, Hao Bai, Pascal Briois, and Fei Gao. 2019. Data-Fusion Prognostics of Proton Exchange Membrane Fuel Cell Degradation. *IEEE Transactions on Industry Applications* 55: 4321–31. [CrossRef]
- McNally, Sean, Jason Roche, and Simon Caton. 2018. Predicting the Price of Bitcoin Using Machine Learning. Paper presented at THE 2018 26th Euromicro International Conference on Parallel, Distributed and Network-Based Processing (PDP), Cambridge, UK, March 21–23.
- Morlet, Jetal, G. Arens, Eliane Fourgeau, and D. Giard. 1982. Wave Propagation and Sampling Theory—Part II: Sampling Theory and Complex Waves. *Geophysics* 47: 222–36. [CrossRef]
- Moustafa, Sayed S. R., and Sara S. Khodairy. 2023. Comparison of different predictive models and their effectiveness in sunspot number prediction. *Physica Scripta* 98: 045022. [CrossRef]
- Mukta, Md Saddam Hossain, Md Adnanul Islam, Faisal Ahamed Khan, Afjal Hossain, Shuvanon Razik, Shazzad Hossain, and Jalal Mahmud. 2022. A Comprehensive Guideline for Bengali Sentiment Annotation. *ACM Transactions on Asian and Low-Resource Language Information Processing* 21: 1–19. [CrossRef]
- Mussa, Michael. 1976. The Exchange Rate, the Balance of Payments and Monetary and Fiscal Policy under a Regime of Controlled Floating. *The Scandinavian Journal of Economics* 78: 229. [CrossRef]
- Naeem, Samreen, Wali Khan Mashwani, Aqib Ali, M. Irfan Uddin, Marwan Mahmoud, Farrukh Jamal, and Christophe Chesneau. 2021. Machine Learning-Based USD/PKR Exchange Rate Forecasting Using Sentiment Analysis of Twitter Data. *Computers, Materials & Continua* 67: 3451–61. [CrossRef]
- Nguyen, Thi Thu Giang, and Robert Ślepaczuk. 2022. *The Efficiency of Various Types of Input Layers of LSTM Model in Investment Strategies on S&P500 Index*. (No. 2022-29). St. Louis: Research Papers in Economics.
- Nwosu, Ugochinyere Ihuoma, Chukwudi Paul Obite, Prince Henry Osuagwu, and Obioma Gertrude Onukwube. 2021. Modeling the British Pound Sterling to Nigerian Naira Exchange Rate During the COVID-19 Pandemic. *Journal of Mathematics and Statistics Studies* 2: 25–35. [CrossRef]
- Pai, Ping-Feng, and Chia-Hsin Liu. 2018. Predicting Vehicle Sales by Sentiment Analysis of Twitter Data and Stock Market Values. *IEEE Access* 6: 57655–62. [CrossRef]
- Philander, Kahlil, and YunYing Zhong. 2016. Twitter Sentiment Analysis: Capturing Sentiment from Integrated Resort Tweets. *International Journal of Hospitality Management* 55: 16–24. [CrossRef]
- Phyo, Pyae-Pyae, Yung-Cheol Byun, and Namje Park. 2022. Short-Term Energy Forecasting Using Machine-Learning-Based Ensemble Voting Regression. *Symmetry* 14: 160. [CrossRef]
- Plakandaras, Vasilios, Theophilos Papadimitriou, and Periklis Gogas. 2015. Forecasting Daily and Monthly Exchange Rates with Machine Learning Techniques. *Journal of Forecasting* 34: 560–73. [CrossRef]
- Qiu, Yue, Zhewei Song, and Zhensong Chen. 2022. Short-Term Stock Trends Prediction Based on Sentiment Analysis and Machine Learning. *Soft Computing* 26: 2209–24. [CrossRef]
- Ratner, Mitchell, and Steven Klein. 2008. The Portfolio Implications of Gold Investment. *The Journal of Investing* 17: 77–87. [CrossRef]
- Raza, Syed Ali, Nida Shah, and Muhammad Shahbaz. 2018. Does Economic Policy Uncertainty Influence Gold Prices? Evidence from a Nonparametric Causality-In-Quantiles Approach. *Resources Policy* 57: 61–68. [CrossRef]
- Razzaq, Abdul, Muhammad Asim, Zulqrnain Ali, Salman Qadri, Imran Mumtaz, Dost Muhammad Khan, and Qasim Niaz. 2019. Text Sentiment Analysis Using Frequency-Based Vigorous Features. *China Communications* 16: 145–53. [CrossRef]

- Razzaque, Mohammad A., Sayema Haque Bidisha, and Bazlul Haque Khondker. 2017. Exchange Rate and Economic Growth. *Journal of South Asian Development* 12: 42–64. [CrossRef]
- Ribeiro, Andrea Maria N. C., Pedro Rafael X. do Carmo, Iago Richard Rodrigues, Djamel Sadok, Theo Lynn, and Patricia Takako Endo. 2020. Short-Term Firm-Level Energy-Consumption Forecasting for Energy-Intensive Manufacturing: A Comparison of Machine Learning and Deep Learning Models. *Algorithms* 13: 274. [CrossRef]
- Sadefo Kamdem, Jules, Rose Bandolo Essomba, and James Njong Berinyuy. 2020. Deep Learning Models for Forecasting and Analyzing the Implications of COVID-19 Spread on Some Commodities Markets Volatilities. *Chaos, Solitons & Fractals* 140: 110215. [CrossRef]
- Seals, Ethan, and Steven R. Price. 2020. Preliminary Investigation in the Use of Sentiment Analysis in Prediction of Stock Forecasting Using Machine Learning. Paper presented at 2020 SoutheastCon, Raleigh, NC, USA, March 28–29.
- Selvin, Sreelekshmy, R. Vinayakumar, E. A. Gopalakrishnan, Vijay Krishna Menon, and K. P. Soman. 2017. Stock Price Prediction Using LSTM, RNN and CNN-Sliding Window Model. Paper presented at the 2017 International Conference on Advances in Computing, Communications and Informatics (ICACCI), Manipal, India, September 13–16.
- Sharma, Urvashi, Rattan K. Datta, and Kavita Pabreja. 2020. Sentiment Analysis and Prediction of Election Results 2018. In *Social Networking and Computational Intelligence*. Berlin: Springer, pp. 727–39. [CrossRef]
- Shih, Han, and Suchithra Rajendran. 2019. Comparison of Time Series Methods and Machine Learning Algorithms for Forecasting Taiwan Blood Services Foundation's Blood Supply. *Journal of Healthcare Engineering* 2019: 6123745. [CrossRef]
- Siami-Namini, Sima, Neda Tavakoli, and Akbar Siami Namin. 2018. A Comparison of ARIMA and LSTM in Forecasting Time Series. Paper presented at 2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA), Orlando, FL, USA, December 17–18.
- Siami-Namini, Sima, Neda Tavakoli, and Akbar Siami Namin. 2019. The Performance of LSTM and BiLSTM in Forecasting Time Series. Paper presented at the 2019 IEEE International Conference on Big Data (Big Data), Los Angeles, CA, USA, December 9–12.
- Sivri, Mahmut Sami, Alp Ustundag, and Buse Sibel Korkmaz. 2022. Ensemble Learning Based Stock Market Prediction Enhanced with Sentiment Analysis. Paper presented at the INFUS 2021 Conference, Intelligent and Fuzzy Techniques for Emerging Conditions and Digital Transformation, Izmir, Turkey, August 24–26; vol. 2, pp. 446–54.
- Smailović, Jasmina, Miha Grčar, Nada Lavrač, and Martin Žnidaršič. 2013. Predictive Sentiment Analysis of Tweets: A Stock Market Application. Paper presented at the Human-Computer Interaction and Knowledge Discovery in Complex, Unstructured, Big Data, Maribor, Slovenia, July 1–3; pp. 77–88. [CrossRef]
- Sun, Chen, Yingxiong Nong, Zhibin Chen, Dong Liang, Ying Lu, and Yishuang Qin. 2022. The CEEMD-LSTM-ARIMA Model and Its Application in Time Series Prediction. *Journal of Physics: Conference Series* 2179: 012012. [CrossRef]
- Vijayarani, S., and R. Janani. 2016. Text Mining: Open Source Tokenization Tools—An Analysis. *Advanced Computational Intelligence: An International Journal (ACII)* 3: 37–47. [CrossRef]
- Wang, Jie, and Jun Wang. 2016. Forecasting Energy Market Indices with Recurrent Neural Networks: Case Study of Crude Oil Price Fluctuations. *Energy* 102: 365–74. [CrossRef]
- Wu, Junhao, and Zhaocai Wang. 2022. A Hybrid Model for Water Quality Prediction Based on an Artificial Neural Network, Wavelet Transform, and Long Short-Term Memory. *Water* 14: 610. [CrossRef]
- Wu, Xianghua, Jieqin Zhou, Huaying Yu, Duanyang Liu, Kang Xie, Yiqi Chen, Jingbiao Hu, Haiyan Sun, and Feng-Juan Xing. 2021. The Development of a Hybrid Wavelet-ARIMA-LSTM Model for Precipitation Amounts and Drought Analysis. *Atmosphere* 12: 74. [CrossRef]
- Wysocki, Maciej, and Robert Ślepaczuk. 2022. Artificial Neural Networks Performance in WIG20 Index Options Pricing. *Entropy* 24: 35. [CrossRef] [PubMed]
- Xiang, Nan, Qianqian Jia, and Yuedong Wang. 2021. Sentiment Analysis of Chinese Weibo Combining BERT Model and Hawkes Process. Paper presented at the 2021 5th International Conference on Deep Learning Technologies (ICDLT), Qingdao, China, July 23–25.
- Xue, Sheng, Hualiang Chen, and Xiaoliang Zheng. 2022. Detection and quantification of anomalies in communication networks based on LSTM-ARIMA combined model. *International Journal of Machine Learning and Cybernetics* 13: 3159–72. [CrossRef] [PubMed]
- Yamak, Peter T., Li Yujian, and Pius K. Gadosey. 2019. A Comparison between ARIMA, LSTM, and GRU for Time Series Forecasting. Paper presented at the 2019 2nd International Conference on Algorithms, Computing and Artificial Intelligence, Sanya, China, December 20–22.
- Yu, He, Li Ming, Ruan Sumei, and Zhao Shuping. 2020. A Hybrid Model for Financial Time Series Forecasting—Integration of EWT, ARIMA with the Improved ABC Optimized ELM. *IEEE Access* 8: 84501–18. [CrossRef]
- Zhang, Qiang, Feng Li, Fei Long, and Qiang Ling. 2018. Vehicle Emission Forecasting Based on Wavelet Transform and Long Short-Term Memory Network. *IEEE Access* 6: 56984–94. [CrossRef]
- Zhang, Xiaoyu, Stefanie Kuenzel, Nicolo Colombo, and Chris Watkins. 2022. Hybrid Short-Term Load Forecasting Method Based on Empirical Wavelet Transform and Bidirectional Long Short-Term Memory Neural Networks. *Journal of Modern Power Systems and Clean Energy* 10: 1216–28. [CrossRef]
- Zhang, Yuchen, and Shigeyuki Hamori. 2020. The Predictability of the Exchange Rate When Combining Machine Learning and Fundamental Models. *Journal of Risk and Financial Management* 13: 48. [CrossRef]
- Zhao, Jiwei, Guangzheng Nie, and Yihao Wen. 2022. Monthly Precipitation Prediction in Luoyang City Based on EEMD-LSTM-ARIMA Model. *Water Science and Technology* 87: 318–35. [CrossRef]


Zhou, Yong, Li Wang, and Junhao Qian. 2022. Application of Combined Models Based on Empirical Mode Decomposition, Deep Learning, and Autoregressive Integrated Moving Average Model for Short-Term Heating Load Predictions. *Sustainability* 14: 7349. [CrossRef]

Zolfaghari, Mehdi, and Samad Gholami. 2021. A Hybrid Approach of Adaptive Wavelet Transform, Long Short-Term Memory and ARIMA-GARCH Family Models for the Stock Index Prediction. *Expert Systems with Applications* 182: 115149. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

Article

Pricing Multi-Asset Bermudan Commodity Options with Stochastic Volatility Using Neural Networks

Kentaro Hoshisashi ^{1,2,3} and Yuji Yamada ^{4,*} 

¹ Department of Computer Science, University College London, London WC1E 6BT, UK; kentaro.hoshisashi.22@ucl.ac.uk

² Graduate School of Business Sciences, University of Tsukuba, Tokyo 112-0012, Japan

³ Sumitomo Mitsui Banking Corporation, Tokyo 100-0005, Japan

⁴ Faculty of Business Sciences, University of Tsukuba, Tokyo 112-0012, Japan

* Correspondence: yuji@gssm.otsuka.tsukuba.ac.jp

Abstract: It has been recognized that volatility in commodity markets fluctuates significantly depending on the demand–supply relationship and geopolitical risk, and that risk and financial management using multivariate derivatives are becoming more important. This study illustrates an application of multi-layered neural networks for multi-dimensional Bermudan option pricing problems assuming a multi-asset stochastic volatility model in commodity markets. In addition, we aim to identify continuation value functions for these option pricing problems by implementing smooth activation functions in the neural networks and evaluating their accuracy compared with other activation functions or regression techniques. First, we express the underlying asset dynamics using the multi-asset stochastic volatility model with mean reversion properties in the commodity market and formulate the multivariate Bermudan commodity option pricing problem. Subsequently, we apply multi-layer perceptrons in the neural network to represent the continuation value functions of Bermudan commodity options, wherein the entire neural network is trained using the least-squares Monte Carlo simulation method. Finally, we perform numerical experiments and demonstrate that applications of neural networks for Bermudan options in a multi-dimensional commodity market achieve sufficient accuracy with regard to various aspects, including changing the exercise dates, the number of layers/neurons, and the dimension of the problem.

Keywords: Bermudan commodity options; multi-layer perceptron; multi-asset stochastic volatility model



Citation: Hoshisashi, Kentaro, and Yuji Yamada. 2023. Pricing Multi-Asset Bermudan Commodity Options with Stochastic Volatility Using Neural Networks. *Journal of Risk and Financial Management* 16: 192. <https://doi.org/10.3390/jrfm16030192>

Academic Editors: Kentaro Iwatsubo and Thanasis Stengos

Received: 28 December 2022

Revised: 5 March 2023

Accepted: 10 March 2023

Published: 12 March 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

In commodity markets, typical products include directional trades such as futures and forwards, which establish an obligation to purchase or sell an underlying commodity in the future (Clark 2014). As essential tools for managing risks from these contracts, which may consist of multiple underlying assets, there are various options contracts that provide a right to trade the underlying commodity under a specified condition. In this study, we focus on early-exercisable options on multiple underlying assets in commodity markets, i.e., multivariate Bermudan commodity options.

Solving Bermudan commodity option pricing problems with multiple underlying assets and factors is challenging because computational efforts grow exponentially in tandem with the problem dimension in general, which is determined by the number of assets and factors. However, the improvement of algorithms and the rapid growth of computational power have led to a remarkable surge of interest in computational science in recent years. Currently, a wide variety of machine learning algorithms, such as deep learning and neural networks, are successfully employed for classification, regression, clustering, or dimensionality reduction tasks and are applied for large-sized and high-dimensional data in various areas. In this study, we develop a new Bermudan commodity option algorithm

via multi-layered neural networks and show its efficiency and effectiveness based on the multi-asset commodity market model with stochastic volatility, wherein significant changes in volatility may be observed according to the demand–supply relationship and geopolitical conditions in commodity markets.

An important feature of Bermudan options is that they can be exercised early, with their value being determined by whether or not they are exercised before maturity. In other words, the option holder must decide whether to continue holding the option or immediately exercise it at a prespecified period. In this situation, it is crucial to determine the continuation value, i.e., the value of holding the option until the next exercise window. Such a continuation value may be given as the discounted conditional expectation of the remaining option value on one step ahead under a risk-neutral probability measure, which generally has no explicit boundary conditions. Moreover, the conditional expectation providing the continuation value is an unknown (possibly nonlinear and complex) multivariate function whose dimensions depend on the number of underlying factors; hence, an exact (yet still approximate) computation involves high-dimensional discrete grids concerning state variables and is quite difficult to solve.

To price early-exercisable options with estimations of continuation value functions, Longstaff and Schwartz (2001) proposed a simple yet powerful numerical method involving a regression-based functional estimation using simulated sample paths known as the least-squares Monte Carlo (LSMC) method. Since then, several studies have examined the application of neural networks (or machine learning methods) for estimating continuation value functions in option pricing based on the LSMC method. For American options' pricing, Haugh and Kogan (2004) applied a neural network with one hidden layer for valuation, whereas Kohler et al. (2010) proved price consistency and convergence with multiple payoff types. Lapeyre and Lelong (2021) gave several numerical examples of Bermudan options and proved convergence. There are additional examples, e.g., the 5000 assets rainbow option (Becker et al. 2021) and expected exposures (Andersson and Oosterlee 2021). Furthermore, other machine learning algorithms have been used for early-exercisable options, e.g., radial basis functions (Ballestra and Pacelli 2013), nearest-neighbor (Feng et al. 2013), deep learning (Becker et al. 2020; Liang et al. 2021), unsupervised learning (Salvador et al. 2020), and reinforcement learning (Li 2022), as well as the support vector machine (Lin and Almeida 2021). Moreover, numerical approaches have also been used, including stochastic kriging metamodels (Ludkovski 2018), high-dimensional partial differential equations (Sirignano and Spiliopoulos 2018), and backward stochastic differential equations (Chen and Wan 2021). Furthermore, there are other applications of neural networks in the finance field, e.g., extending the feature set (Montesdeoca and Niranjana 2016), the calculation of implied volatilities (Liu et al. 2019), and decision-making (Puka et al. 2021). A comprehensive review of these methods was conducted by Ruf and Wang (2020).

Although this study shares the same ideas as the aforementioned studies—in particular, as in Lapeyre and Lelong (2021), given that a multi-layer perceptron (MLP) is applied in the neural network—it is worthwhile to mention that our study may be considered novel in several aspects: We illustrate an algorithm for estimating the continuation values of multi-asset Bermudan commodity options with stochastic volatility features, whereby a smooth activation function, such as the sigmoid function, is applied in the MLP to reflect the smoothness of conditional expectations regarding state variables. The smoothness of functions to represent conditional expectations is key in this study. In the Markovian setting, the conditional probability density functions are usually smooth given state variables; thus, conditional expectations are smooth functions. Therefore, the target continuation value function is smooth, and we can expect a better fit to the target function by using a smooth activation function in the MLP. This is in contrast with more commonly used piecewise linear functions such as the leaky ReLU function applied in the numerical experiments by Lapeyre and Lelong (2021), wherein the fitting function may not be smooth but only piecewise smooth. (Note that, in a similar context with the smoothness of estimated functions, Yamada (2012, 2017) applied the generalized additive model to calculate smooth

functions for conditional expectations in multivariate hedging problems with European and Bermudan options.) Additionally, we applied more sophisticated techniques such as the resampling procedure and early stopping to improve the computational efficiency and avoid possible biases in pricing or overfitting for optimal learning (see Section 3.2 and the numerical experiments).

While several neural network/machine learning models for option pricing exist, we believe that, for this type of research, the current methodologies, along with developed computational algorithms, need to be combined with existing techniques using the currently available computational environment. In this context, the choice of problem and methodology choice is important, as is the approach to the problem and how to perform the numerical experiments. The combination of multi-asset commodity options with stochastic volatility and recently developed neural network techniques (including the computational environment and software) is meaningful since commodity markets are largely volatile, and this volatility may change over time. Moreover, the multi-asset Bermudan commodity options with stochastic volatility, to the best of our knowledge, have not been previously considered despite the problem's importance. It should become more challenging in numerical calculations to configure multiple underlying assets that recognize mean-reverting dynamics and solve boundary conditions with stochastic factors (see, e.g., Hahn and Dyer 2008 and Ball and Roma 1994).

The present study implements a multi-layered neural network and examines its efficiency and effectiveness for multi-asset Bermudan commodity option pricing problems with stochastic volatility. First, we formulate the multi-asset commodity market with stochastic volatility, wherein individual asset price dynamics are expressed as a two-factor model by combining a well-known commodity model by Schwartz (1997) with Heston's stochastic volatility model (Heston 1993). Next, we apply MLPs in the neural network to represent the continuation value functions in Bermudan option pricing, whereby the entire neural network is trained using LSMC simulations. We perform numerical experiments to compare the continuation value function accuracy in response to changing the exercise dates, the number of layers/neurons, and the dimension of the problem. We also compare the relationship between the continuation values and network configurations.

The outline of this article is as follows. Section 2 gives an introduction to the commodity option structure adopted in this study and the formulation of the multi-dimensional Bermudan option problem. Section 3 describes the configuration of neural networks, a multi-dimensional asset model with stochastic volatility, and a Bermudan options pricing procedure for learning and valuation via Monte Carlo sample paths. Section 4 presents the numerical results of the Bermudan option prices and compares the accuracy of the continuation value surfaces. Section 5 summarizes the analysis results and discussions. Lastly, Section 6 concludes this study.

2. Pricing Multi-Asset Bermudan Commodity Options with Stochastic Volatility

In this section, we introduce early-exercisable commodity options and formulate the problem of pricing multi-asset Bermudan commodity options with stochastic volatility.

2.1. Early-Exercisable Commodity Options

As stated earlier, in commodity markets, typical products include plain directional trades such as future and forward contracts, which establish an obligation to buy or sell a particular commodity asset at a specified price in the future (Clark 2014). Depending on the terminal values of commodity assets, holding these contracts may lead to a loss or profit for the contractor, while a large loss is particularly undesirable for the holder; furthermore, the possibility of a large profit may be pursued. Such opportunities are realized using options contracts, giving a right to purchase or sell an underlying commodity asset with a prespecified strike price in the future.

Among the many types of options used as hedging tools in commodity markets, early-exercisable options provide additional flexibility regarding exercise timing and are

considered useful for hedgers, in practice. Traditionally, such options are characterized as American options; however, in the context of exotic options, Bermudan options have similar flexibility. These options allow holders to exercise them early, although only on specific dates before maturity; thus, the option holder must decide whether to continue holding the option or immediately exercise it during the exercisable period. However, such continuation value is usually unknown because it depends on future option values on specific exercisable dates. Therefore, it is paramount to determine the continuation values for Bermudan options. The objective of this study is to evaluate computational performance (including the accuracy of continuation value estimation) for pricing multi-asset Bermudan commodity options via multi-layered neural networks.

2.2. Multi-Dimensional Bermudan Option Pricing Problem

In this subsection, we describe the multi-dimensional Bermudan option pricing problem, following Lapeyre and Lelong (2021). Given a complete filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{0 \leq t \leq T}, \mathbb{P})$ with a finite time horizon $T > 0$, we assume that a set of underlying assets is modeled via a multifactored process $(X_t)_{0 \leq t \leq T}$ adapted to the filtration, $(\mathcal{F}_t)_{0 \leq t \leq T}$, and that \mathbb{P} is an associated risk-neutral measure. We consider a Bermudan option with exercise dates $0 = T_0 \leq T_1 < T_2 < \dots < T_N = T$ and a discrete-time payoff process P_{T_n} if exercised at times $(T_n)_{0 \leq n \leq N}$, where P_{T_n} is specified as a function of X_{T_n} . Then, Bermudan option prices Z_{T_n} are computed using the following recursive equation:

$$\begin{cases} Z_{T_N} = P_{T_N} \\ Z_{T_n} = \max(P_{T_n}, e^{-r\delta_{T_n}} \mathbb{E}[Z_{T_{n+1}} | \mathcal{F}_{T_n}]) \end{cases}, 0 \leq n \leq N - 1 \tag{1}$$

where \mathbb{E} denotes the expectation under the risk-neutral probability measure \mathbb{P} with the risk-free interest rate r and the interval between T_{n-1} and T_n as δ_{T_n} . Furthermore, assuming that $(X_t)_{0 \leq t \leq T}$ is a multi-dimensional Markov process, there exists a measurable function $\Phi_n : \mathbb{R}^{d_x} \rightarrow \mathbb{R}$, such that:

$$e^{-r\delta_{T_n}} \mathbb{E}[Z_{T_{n+1}} | \mathcal{F}_{T_n}] = e^{-r\delta_{T_n}} \mathbb{E}[Z_{T_{n+1}} | X_{T_n}] = \Phi_n(X_{T_n}), 0 \leq n \leq N - 1. \tag{2}$$

Herein, we refer to Φ_n as a continuation value function in this paper.

Note that finding the exact Φ_n is difficult; alternatively, one may identify a function f_n to minimize the following quantity,

$$\mathbb{E} \left[\left| e^{-r\delta_{T_n}} Z_{T_{n+1}} - f_n(X_{T_n}) \right|^2 \right], \tag{3}$$

over a parametrized set of functions \mathfrak{B} . If all (real-valued) square-integrable measurable functions are searched to minimize Equation (3), it turns out that the function Φ_n providing the conditional expectation in Equation (2) is achieved via an optimizer. However, there is a trade-off between the generality of a set of functions and the efficiency of computation. Additionally, computational tractability depends on the methodology to solve the optimization problem.

2.3. Multi-Asset Commodity Market Model with Stochastic Volatility

This study employs a multivariate commodity market model consisting of multiple underlying assets with stochastic volatility for the Bermudan option problem. To this end, we adopt a stochastic volatility model for the mean-reverting commodity dynamics (Schwartz 1997) and expand it to the multi-asset case.

Consider the Bermudan option problem with n underlying assets at time t , $S_{i,t}$, $i = 1, \dots, l$, the i -th price dynamics of which are governed by the following two-dimensional stochastic differential equations (SDEs):

$$\begin{aligned} dS_{i,t} &= \kappa_{S_i}(\mu_i - \ln S_{i,t})S_{i,t}dt + \sqrt{v_{i,t}}S_{i,t} dW_{S_{i,t}}, \\ dv_{i,t} &= \kappa_{v_i}(\theta_i - v_{i,t})dt + \zeta_i \sqrt{v_{i,t}}dW_{v_{i,t}}. \end{aligned} \tag{4}$$

Herein, $W_{S_i,t}$ and $W_{v_i,t}$ are correlated Brownian motions with appropriate correlation parameters, while the magnitude of the speed coefficient κ_{S_i} measures the degree of mean reversion to the long-run mean μ_i , including the market price of risk in the underlying asset price processes. The second term characterizes the i -th volatility process, $\sigma_{i,t} \equiv \sqrt{v_{i,t}}$, where κ_{v_i} indicates a degree of mean reversion toward long-term volatility θ_i , and ξ_i is the volatility of volatility.

Since each of the underlying asset price dynamics in (4) follows a two-dimensional Markov process, the state variables at time t , denoted by X_t , corresponding to the input features X of the MLP in the previous subsection, may be described as

$$X_t := [S_{1,t}, \sigma_{1,t}, S_{2,t}, \sigma_{2,t}, \dots, S_{l,t}, \sigma_{l,t}]^\top \in \mathbb{R}^{d_l}. \tag{5}$$

The dimension d_l in (5) depends on the number of state variables and is given by $d_l := 2l$. Note that the SDEs of the underlying assets are used to generate sample paths of the LSMC method in the Bermudan option pricing problem.

3. Application of Neural Networks with MLP

When pricing Bermudan commodity options using a model with multi-dimensional factors—including the multi-asset stochastic volatility model introduced in the previous section—it is crucial to determine continuation values at each exercisable date. In order to identify a continuation value function in the multi-asset Bermudan option pricing problem with stochastic volatility, this study takes a neural network approach with MLP, similar to Lapeyre and Lelong (2021). First, we introduce the neural network architecture considered in this study, which generates a continuation value in Bermudan commodity options pricing. Second, we present the underlying assets model with multi-dimensional factors, which has multi-asset and stochastic volatility. Finally, we provide algorithms for learning the entire network and option pricing procedure.

3.1. Continuation Value Functions via MLPs

First, we explain the configuration of an MLP to express a general multi-dimensional function and to approximate the continuation value function in the multi-dimensional Bermudan option problem.

The basic configuration of the MLP is shown in Figure 1, where $X \in \mathbb{R}^d$ is an input vector and $Z \in \mathbb{R}$ is an output of the entire neural network. Each neuron is called a “perceptron” that defines a mapping of input/output signals with appropriate dimensions, being dependent on the number of neurons at input/output layers. For example, if $x \in \mathbb{R}^{d_x}$ denotes an input signal of a perceptron with a weight matrix $W \in \mathbb{R}^{d_x \times d_y}$ and a bias vector $b \in \mathbb{R}^{d_y}$, then, an output signal $y \in \mathbb{R}^{d_y}$ from the perceptron is given by

$$y = h(W^T x + b), \tag{6}$$

where $h : \mathbb{R}^{d_y} \rightarrow \mathbb{R}^{d_y}$ is a component-wise activation function. Typical choices of activation functions in neurons are as follows:

$$\begin{aligned} \text{Sigmoid} : x &\mapsto \frac{1}{1+e^{-x}} \\ \text{ReLU} : x &\mapsto \max(x, 0) \end{aligned} \tag{7}$$

In the case where the MLP is applied for a regression, all the weight matrices and bias vectors in the MLP are computed to minimize the sum of squared errors between the actual dependent variable, denoted by $Z \in \mathbb{R}$, and the predicted dependent variable $\hat{Z} \in \mathbb{R}$ given the training datasets of $X \in \mathbb{R}^{d_x}$ and $Z \in \mathbb{R}$ (expressed using the MLP in Figure 1).

Note that the MLP can express a continuous and complex nonlinear surface in entire networks by sequentially performing a linear and nonlinear transformation on inputs $X \in \mathbb{R}^{d_x}$ to the compiled layer output $\hat{Z} \in \mathbb{R}$. The properties of the MLP function derive from the universal approximation theorem proposed by Cybenko (1989) and the Kolmogorov–

Arnold representation theorem put forward by Kolmogorov (1957) and Arnold (2009), in which any function can be approximated if the input size and network are infinite. In this sense, functions expressed by the MLP are generally considered suitable for a problem with complicated interactions because of the adjustable basis functions (see Choon et al. 2008).

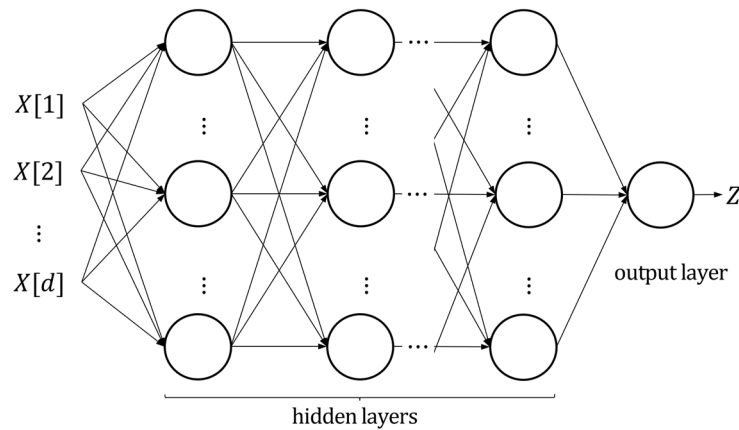


Figure 1. Multi-layered structure with perceptrons (circles) in a multi-layer perceptron.

For the randomly generated sample paths of $(X_t)_{0 \leq t \leq T}$, we can apply the LSMC method (see Appendices A and B) combined with the MLP, whereby the continuation value function is modeled at each step using a function given by the MLP. Let $X_t^{(1)}, X_t^{(2)}, \dots, X_t^{(M)}$ and $0 \leq t \leq T$ be the simulated M sample paths of $(X_t)_{0 \leq t \leq T}$. Since the discrete-time payoff process P_{T_n} (if exercised at times $(T_n)_{0 \leq n \leq N}$) is specified as a function of X_{T_n} for a Bermudan option and $Z_{T_N} = P_{T_N}$, the training data of the output variable, $Z \equiv Z_{T_N} \in \mathbb{R}$, in the first step of the LSMC method, are computed as $P_{T_N}^{(1)}, P_{T_N}^{(2)}, \dots, P_{T_N}^{(M)}$, corresponding to the payoffs of the Bermudan option at maturity along the sample path of X_{T_N} . The MLP in the first step is constructed for the training data of $X \equiv X_{T_{N-1}} \in \mathbb{R}^{d_x}$, given as $X_{T_{N-1}}^{(1)}, X_{T_{N-1}}^{(2)}, \dots, X_{T_{N-1}}^{(M)}$, together with those of $Z \equiv Z_{T_N} \in \mathbb{R}$. Then, we obtain an approximation of the continuation value function, denoted by $\hat{\Phi}_{N-1}$, and the continuation values along the sample path, $\hat{\Phi}_{N-1}(X_{T_{N-1}}^{(m)})$, $m = 1, \dots, M$.

In the second step, the training data of the output variable, $Z \equiv Z_{T_{N-1}} \in \mathbb{R}$, are computed using (1), as

$$Z_{T_{N-1}}^{(m)} = \max(P_{T_{N-1}}^{(m)}, \hat{\Phi}_{N-1}(X_{T_{N-1}}^{(m)})), \quad m = 1, \dots, M, \tag{8}$$

as well as the training data of $X \equiv X_{T_{N-2}} \in \mathbb{R}^{d_x}$, given as $X_{T_{N-2}}^{(1)}, X_{T_{N-2}}^{(2)}, \dots, X_{T_{N-2}}^{(M)}$. Using these training datasets, the MLP is constructed to find an approximation of the continuation value function, denoted by $\hat{\Phi}_{N-2}$, and the continuation values along the sample paths, $\hat{\Phi}_{N-2}(X_{T_{N-2}}^{(m)})$. We then repeat the same procedure until T_0 .

3.2. Learning Networks and Option Pricing

For learning neural networks, we generate Monte Carlo sample paths using the SDEs in Section 3.2 based on a similar idea to that of the ordinary LSMC method introduced by Longstaff and Schwartz (2001). Herein, we apply nonlinear functions of the MLPs instead of polynomial functions for the basis of the continuation value functions. Additionally, we introduce techniques such as early stopping to improve the fitted continuation functions and avoid possible overfitting or biases in learning and pricing. We also introduce the resampling procedure to avoid a possible bias caused by using the same random samples between learning and valuation, and we regenerate Monte Carlo sample paths for the valuation of Bermudan option prices (see Appendix A for pricing details).

Herein, we summarize a learning procedure, as described in Algorithm 1 below, where the underlying price and the volatility vector are denoted by $S_t := [S_{1,t}, S_{2,t}, \dots, S_{l,t}]^\top$ and $\sigma_t := [\sigma_{1,t}, \sigma_{2,t}, \dots, \sigma_{l,t}]^\top$, while g denotes a payoff function of S_t . Given simulation the sample paths generated by the multi-asset stochastic volatility models in (4), we provide a Bermudan option pricing procedure involving the algorithm for estimating the continuation values using neural networks. This algorithm operates to find MLP Φ as a continuation value function satisfying Equation (8) in the previous section and gives the Bermudan option price.

Algorithm 1. Bermudan option pricing with learning networks.

Require: Initiate paths $S_t^{(j)}, \sigma_t^{(j)}, t = T_0, T_1, \dots, T_N, j = 1, 2, \dots, M$
1: Let p be the patience and Max_{iter} be the maximum number of epochs
2: Put $V^{(j)} \leftarrow g(S_{T_N}^{(j)})$ for all j
3: **for** t from T_{N-1} to T_1 **do**
4: Let $X^{(j)} \leftarrow S_t^{(j)}, \sigma_t^{(j)}$ and $V^{(j)} \leftarrow e^{-r\delta t} \cdot V^{(j)}$ for all j
5: **if** t on exercisable periods **then**
6: Perform learning on X to obtain network Φ_t with Z to be V
7: $i \leftarrow 0$
8: $k \leftarrow 0$
9: **while** $i < Max_{iter}$ **do**
10: Train Φ_t on X and V
11: **if** improved **then**
12: $k \leftarrow 0$
13: **else**
14: $k \leftarrow k + 1$
15: **end if**
16: **if** $k == p$ **then**
17: Break
18: **end if**
19: $i \leftarrow i + 1$
20: **end while**
21: Calculate the continuation value $\Phi_t(X^{(j)})$ for all j
22: **for** j from 1 to M **do**
23: **if** $g(S_t^{(j)}) > \Phi_t(X^{(j)})$ **then**
24: $V^{(j)} \leftarrow g(S_t^{(j)})$
25: **end if**
26: **end for**
27: **end if**
28: **end for**
29: **return** mean of $e^{-r\delta t} \cdot V$

^{Note} This study does not use the selection technique, which performs regression using only the in-the-money paths proposed by Longstaff and Schwartz (2001), for the purpose of constructing a versatile algorithm.

It is noted that one cycle of training with the complete training data is known as an epoch and is repeated for learning purposes for each continuation value function in Algorithm 1. In general, the larger the number of epochs, the better learning of the training data. However, a large number of epochs usually requires a long computational time, even with large computer resources, and sometimes leads to overfitting of the training data. To prevent such situations, we introduce an early stopping rule for the learning procedure given a specified integer p in Algorithm 1. Under the early stopping rule, the objective function (3) is monitored for improvement, and the number of iterations (i.e., the number of epochs) without improvement (compared with the previous epoch), denoted by k , are counted. If this number reaches p , the iteration stops and the learning procedure of the continuation value function terminates; otherwise, the iteration continues as long as the

iteration index i is less than Max_{iter} , where Max_{iter} is the maximum number of epochs specified at the beginning of Algorithm 1. Note that the introduction of the early stopping rule not only decreases the computational time but also prevents overfitting/underfitting for the MLP.

Based on the network configuration of the neural networks, the computational complexity of Algorithm 1 is given by the number of iterations for parameter estimation of the MLP. This number of iterations depends on the maximum number of epochs and the number of exercisable dates, Max_{iter} and $N - 1$. Once these values are specified, the maximum number of iterations is $Max_{iter} \times (N - 1)$, which is the total number of epochs applied in Algorithm 1. In addition, the computational complexity of each epoch in the MLP depends on the network configuration (see Serpen and Gao 2014).

To price the Bermudan commodity options using the continuation functions estimated in Algorithm 1, we regenerate different sample paths from those used in the learning procedure for computing the continuation values and Bermudan commodity option prices, given the neural networks in Algorithm 1, i.e., we separate the learning and the valuation procedures, and Algorithm 1 may be applied without learning (i.e., given the estimated neural networks) for the valuation procedure. The merit of this resampling is that it avoids a price bias, which results from overfitting using the same sample paths. Accordingly, this study adopts the following procedure:

1. Generate the sample paths of the underlying assets for the MLP learning of Algorithm 1.
2. Find the MLP network parameters via learning in Algorithm 1 using the sample paths in Step 1.
3. Given the estimated neural networks, regenerate a different set of sample paths and apply Algorithm 1 (without learning) to compute the continuation values and the initial prices of the Bermudan options.
4. Repeat Step 2 and calculate statistical values such as the mean and the standard deviation of the Bermudan option prices.

In the above, it is key that learning and pricing (i.e., valuation) utilize the different simulation sample paths set in Steps 1 and 3. Figure 2 shows a flowchart of the entire procedure for learning and valuation.

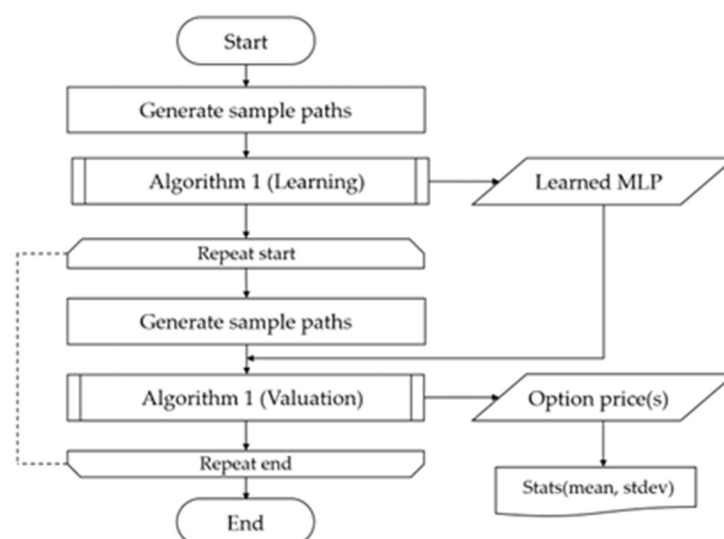


Figure 2. Flowchart of the entire procedure for learning and valuation of Bermudan option pricing.

4. Numerical Experiments

The objective of this section is to execute numerical experiments based on the learning and valuation procedure explained in the previous section and make comparisons of the Bermudan option pricing between the MLP and the benchmark polynomial regression (i.e., the standard (naïve) LSMC method by Longstaff and Schwartz 2001).

4.1. Problem Setting and Preliminary Experiment

Herein, we consider Bermudan commodity options with early-exercisable dates in discretized periods until maturity T , i.e., $0 = T_0 < T_1 < \dots < T_N = T$, the payoffs of which are given by $g(S_t)$ when exercised. We define several settings for different dimensions of Bermudan options, d_l (i.e., the number of state variables in (5)), exercisable dates, and payoff functions. We also introduce a constant volatility model as a one-dimensional problem to perform a preliminary experiment.

For the exercisable dates of the Bermudan commodity options, we consider two cases as depicted in Figure 3. One is a two-period problem, wherein the Bermudan commodity option is issued at time T_0 and can be exercised at T_1 and maturity T_2 . The other is a case with multiple exercisable dates, wherein we choose ten exercisable timings before maturity. In both cases, the options can be exercised after a half-year period to compare the continuation value surfaces at time T_1 between different methodologies, and the values of the options are evaluated at the initial time period, T_0 .

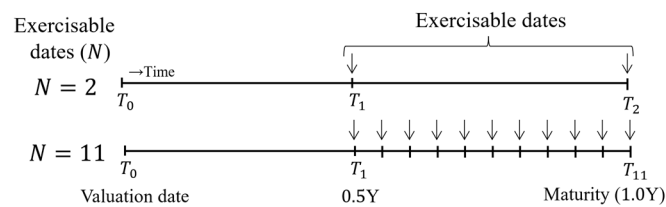


Figure 3. Bermudan option schedules with exercisable dates of 2 and 11 times.

Moreover, the payoff functions for Bermudan basket put options are given as

$$g(S_t) = \max\left(K - \frac{1}{l} \sum_i S_{i,t}, 0\right). \tag{9}$$

Then, an upper limit price D is introduced to the payoff function as

$$g(S_t) = \min\left(\max\left(K - \frac{1}{l} \sum_i S_{i,t}, 0\right), D\right) \tag{10}$$

for Bermudan capped put options. Note that the upper limit, D , provides an additional complexity of the payoff functions.

The parameters of the underlying assets and neural networks are set as shown in Tables 1 and 2 below.

4.2. Low-Dimensional Case: Single-Asset Bermudan Options with Stochastic Volatility and Constant Volatility

We begin with the simplest valuation based on Schwartz’s (1997) single-asset and constant volatility model in, compared with the standard LSMC method using polynomial regression and the finite difference method (FDM) detailed by Tavella and Randall (2000). In the MLP, we applied Algorithm 1 for learning neural networks with Monte Carlo simulations and then repeated the valuation procedure (using Algorithm 1 without learning) 100 times. The MLP in this experiment contains two hidden layers with sixty-four neurons of sigmoid activate functions. In the standard LSMC method, we use a quintic function for one-dimensional problems (i.e., $d_l = 1$) and a multi-dimensional quadratic function for two or more higher-dimensional problems (i.e., $d_l \geq 2$) and apply the same procedure (i.e., the learning and valuation procedure). The FDM is based on the Crank–Nicolson scheme with discretized 2000/200/50 grids in the time/asset/volatility directions.

Table 1. Parameters of the underlying asset model with constant and stochastic volatility.

Parameter	Constant Vol. ($d_t=1$)	Stochastic Vol. ($d_t \geq 2$) ¹
Spot rate (S_0)		100.0
Strike rate (K)		105.0
Capped rate (D)		10.0
Time to maturity (T) [years]		1.0
Risk-free interest rate (r) [%]		6.0
Initial volatility ($\sigma_{i,0}$) [%]		30.0
Long-run mean (μ)		4.8
Kappa of asset (κ_{S_i})		0.3
Long-term volatility ($\sqrt{\theta_i}$) [%]	-	30.0
Correlation (ρ_{S_i, v_i})	-	-0.1
Corr. among assets (ρ_{S_i, S_j})	-	0.7 (1.0 if $i = j$)
Corr. assets and vol. (ρ_{S_i, v_j})	-	-0.07
Corr. among vols. (ρ_{v_i, v_j})	-	0.007 (1.0 if $i = j$)
Kappa of vol. (κ_{v_i})	-	1.5
Vol. of vol. (ζ_i)	-	0.2
Num. of paths (learning)	100,000	
Num. of paths (valuation)	10,000	
Sim. path timesteps (per yr.)	20	

¹ We applied Euler’s method as a discretized method.

Table 2. Neural networks’ learning parameters.

Learning Parameters	Value
Num. of sim. paths (M)	100,000
Batch size	4096
Max. num. of epochs (Max_{iter})	200
Train paths percentage	80%
Evaluation paths percentage	20%
Optimizer	Adam ¹

¹ Learning optimizer Adam (Kingma and Ba 2014) hyperparameters are set to 0.01 for the learning rate, 0.9 for beta1, 0.999 for beta2, and 1×10^{-7} for epsilon; training is completed when the loss does not improve even after 20 epochs, as early stopping. Randomized 20% of input paths are used in evaluations to avoid over-learning.

Table 3 compares the means and standard deviations of the option prices obtained with the MLP and the standard LSMC methods. Considering the option price of the FDM as a proxy for the value of the Bermudan commodity option price, we see that both the MLP and the standard LSMC method almost achieve the Bermudan option price value, i.e., the gap between the three prices is sufficiently small for this one-dimensional problem. We implemented Algorithm 1 using the MLP and the standard LSMC on Python, using a machine learning package based on TensorFlow (Abadi et al. 2015) and the PolynomialFeatures toolbox of the scikit-learn library (Pedregosa et al. 2011). All our numerical experiments were run using Google Colaboratory (Google 2022) with 36 GB of RAM and a dual-core CPU of 2.3 GHz.

Table 3. Bermudan (capped) put option prices in single-asset constant volatility ($d_t = 1$).

Bermudan Put Option ($d_n=1$)				Bermudan Capped Put Option ($d_n=1$)			
	# of Ex.	N = 2	N = 11		# of Ex.	N = 2	N = 11
LSMC price	Mean	11.474	11.786	LSMC price	Mean	5.731	6.244
	(St. dev.)	(0.071)	(0.069)		(St. dev.)	(0.027)	(0.029)
MLP price	Mean	11.471	11.812	MLP price	Mean	5.729	6.320
	(St. dev.)	(0.069)	(0.061)		(St. dev.)	(0.027)	(0.031)
FDM price		11.415	11.808	FDM price		5.752	6.350

^{note} FDM = finite difference method; LSMC = least-squares Monte Carlo; MLP = multi-layer perceptron.

When considering a single-asset stochastic model, the corresponding Bermudan commodity option problem becomes two-dimensional, i.e., $d_I = 2$, we can observe a slight difference between the MLP and the LSMC methods, compared with the approximate solution of the FDM, as shown in Table 4. First, we see that there is no significant difference between the 3 cases for the Bermudan put option with exercisable dates $N = 2$. However, the gap between the LSMC and FDM prices becomes slightly wider, compared with that between the MLP and FDM prices with exercisable dates $N = 11$. For the Bermudan capped put option, there is a slightly larger difference vis-à-vis the FDM price for both the MLP and LSMC prices, whereas a slight improvement was achieved by using the MLP in terms of the gap from the FDM price, as illustrated in the box plots in Figure 4.

Table 4. Bermudan (capped) put option prices in single-asset stochastic volatility ($d_I = 2$).

Bermudan Put Option ($d_I=2$)				Bermudan Capped Put Option ($d_I=2$)			
	# of Ex.	N = 2	N = 11		# of Ex.	N = 2	N = 11
LSMC price	Mean	11.114	11.170	LSMC price	Mean	5.489	5.985
	(St. dev.)	(0.139)	(0.137)		(St. dev.)	(0.047)	(0.043)
MLP price	Mean	11.113	11.417	MLP price	Mean	5.504	6.015
	(St. dev.)	(0.140)	(0.133)		(St. dev.)	(0.047)	(0.045)
FDM price		11.090	11.460	FDM price		5.541	6.070

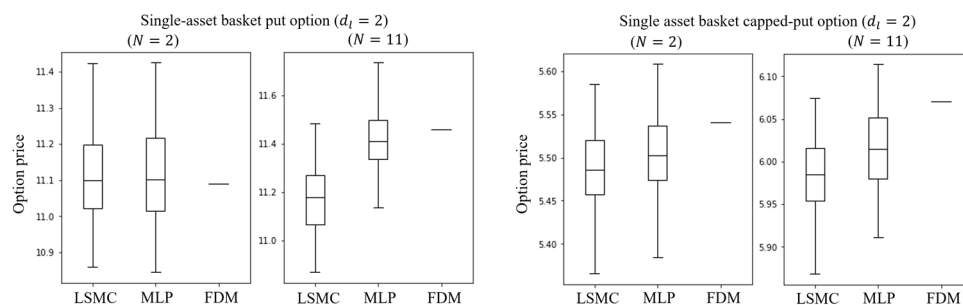


Figure 4. The box plot of option prices with a single asset with stochastic volatility. It has two-dimensional state variables. The conditions were put and capped put payoff, as well as changing exercisable dates N via the method. In the box plot, the center of the box in the first and third quartiles is a median (line), while the beards are the maximum and the minimum values.

4.3. Higher-Dimensional Case: Multi-Asset Bermudan Options with Stochastic Volatility

In the case of higher-dimensional Bermudan commodity options with multi-asset stochastic volatility (e.g., two asset problems with $d_I = 4$), it is difficult (or unrealistic) to obtain an approximate Bermudan option price with high accuracy using the FDM. Thus, we compare option prices obtained with the MLP with the benchmark polynomials (i.e., the standard LSMC method) only, where we set $d_I = 10$, i.e., five-asset stochastic volatility, in the numerical experiments. Then, we will discuss the source of the difference in view of the continuation value functions for both methods in the next section. Additionally, we will compare the accuracy of estimated continuation values by considering a two-asset problem with $d_I = 4$ and two exercisable dates $N = 2$.

Table 5 shows our numerical results, which compare the mean and standard deviation between the LSMC and the MLP prices for $N = 2$ and $N = 11$, obtained via the learning and valuation procedure described in Section 3. In the case of the Bermudan put option for $N = 2$, we see that there is no significant difference between the two methods, as with the case with a low-dimensional problem, $d_I = 2$. However, the gap between the two increased for the Bermudan capped put options and the case with $N = 11$, as shown in the box plots in Figure 5. In other words, we see that the differences between the MLP and the LSMC are emphasized by introducing additional complexity to the payoff function or by increasing exercisable dates.

Table 5. Bermudan (capped) put option prices in the five-asset ($d_t = 10$) stochastic volatility model.

Bermudan Put Option ($d_t=10$)				Bermudan Capped Put Option ($d_t=10$)			
	# of Ex.	N = 2	N = 11		# of Ex.	N = 2	N = 11
LSMC price	Mean	9.712	9.803	LSMC price	Mean	5.406	5.884
	(St. dev.)	(0.108)	(0.105)		(St. dev.)	(0.043)	(0.039)
MLP price	Mean	9.712	9.888	MLP price	Mean	5.414	5.904
	(St. dev.)	(0.103)	(0.098)		(St. dev.)	(0.043)	(0.040)

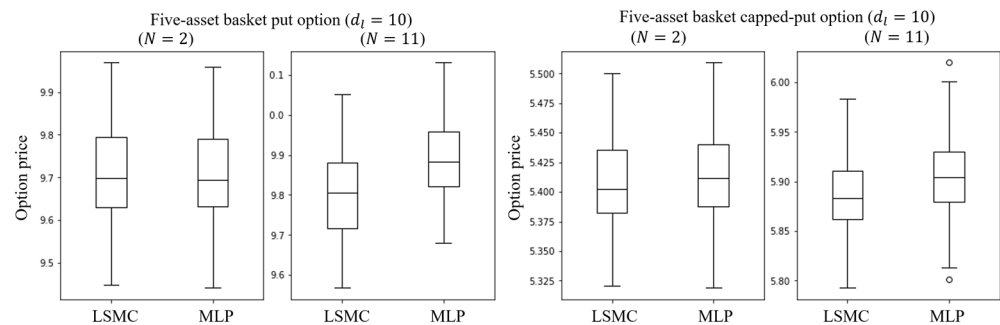


Figure 5. Box plot of option prices’ distribution with five-asset stochastic volatility. It has ten-dimensional state variables. The conditions were put and capped-put payoff, as well as changing exercisable dates N via the method. In the box plot, the circles are outliers.

Additionally, we demonstrated the same numerical experiments above but increased the number of assets in the stochastic volatility model to 10 and 20 assets, respectively (i.e., $d_t = 20$ and $d_t = 40$) and compared the estimated Bermudan commodity option prices between the MLP and the LSMC. To observe the effects of increasing the number of exercisable dates more clearly, we added $N = 6$ between $N = 2$ and $N = 11$ to obtain the estimation values shown in Tables 6 and 7 (corresponding to the cases of $d_t = 20$ and $d_t = 40$, respectively). Table 6 shows the means, the standard deviations, and the gaps between the estimated prices for the Bermudan put options and Bermudan capped put options with $d_t = 20$, and Table 7 shows those with $d_t = 40$.

Table 6. Bermudan (capped) put option prices in the 10-asset ($d_t = 20$) stochastic volatility model.

Bermudan Put Option ($d_t=20$)					Bermudan Capped Put Option ($d_t=20$)				
	# of Ex.	N = 2	N = 6	N = 11		# of Ex.	N = 2	N = 6	N = 11
LSMC price	Mean	9.539	9.595	9.571	LSMC price	Mean	5.359	5.747	5.848
	(St. dev.)	(0.120)	(0.098)	(0.097)		(St. dev.)	(0.036)	(0.042)	(0.041)
MLP price	Mean	9.548	9.681	9.684	MLP price	mean	5.376	5.765	5.851
	(St. dev.)	(0.112)	(0.091)	(0.095)		(St. dev.)	(0.035)	(0.041)	(0.041)
Difference (MLP—LSMC)	Mean	0.009	0.087	0.113	Difference (MLP—LSMC)	Mean	0.016	0.018	0.003

Table 7. Bermudan (capped) put option prices in the 20-asset ($d_t = 40$) stochastic volatility model.

Bermudan Put Option ($d_t=40$)					Bermudan Capped Put Option ($d_t=40$)				
	# of Ex.	N = 2	N = 6	N = 11		# of Ex.	N = 2	N = 6	N = 11
LSMC price	Mean	9.338	9.4855	9.453	LSMC price	Mean	5.362	5.728	5.828
	(St. dev.)	(0.112)	(0.111)	(0.111)		(St. dev.)	(0.048)	(0.042)	(0.037)
MLP price	Mean	9.369	9.596	9.600	MLP price	Mean	5.382	5.750	5.833
	(St. dev.)	(0.111)	(0.105)	(0.104)		(St. dev.)	(0.048)	(0.040)	(0.037)
Difference (MLP—LSMC)	Mean	0.031	0.111	0.147	Difference (MLP—LSMC)	Mean	0.020	0.022	0.005

Similar to the previous cases, the gap between the MLP and LSMC increases for the Bermudan put prices given a larger number of exercisable dates but decreases for the Bermudan capped put options when $N = 11$ for both $d_1 = 20$ and $d_1 = 40$. It is possible that the choice of regression function has a weaker effect for higher dimensional Bermudan capped put options with a larger number of exercisable dates, i.e., the continuation value functions of the capped put options may become flatter or smoother when the number of exercisable dates increases and can be fitted with less sophisticated functions. This phenomenon should be investigated in more detail in a future study.

4.4. Comparison of Continuation Value Surfaces

In the previous subsection, we observed that there are some differences between the MLP and the LSMC regarding the estimated prices and that these differences were more notable for Bermudan capped put options and/or increased exercisable dates. Herein, we discuss the possible reason for this price difference by visualizing the continuation value surfaces for both the MLP and the LSMC methods. For visualization purposes, we consider single-asset Bermudan capped put options with stochastic volatility and constant volatility, i.e., the low-dimensional cases with $d_1 = 1$ and $d_1 = 2$ introduced in Section 4.2.

The left-hand plot of Figure 6 illustrates the continuation value function estimated at T_1 in the one-dimensional Bermudan capped-put option problem (corresponding to the single-asset constant volatility model) using the LSMC, whereas the right-hand plot shows the problem using the MLP. We first observe that the continuation value function of the MLP monotonically decreases with the underlying price, whereas the one obtained from the LSMC method is a nonmonotonic function. Since the payoff function of the Bermudan capped put option is piecewise linear and it monotonically decreases with the underlying price, the monotonicity of the continuation value function is more consistent with the payoff structure of the Bermudan capped-put option. In this sense, we see that the continuation value function of the MLP reflects the monotonic property more appropriately than that of the LSMC method in the one-dimensional single-asset problem.

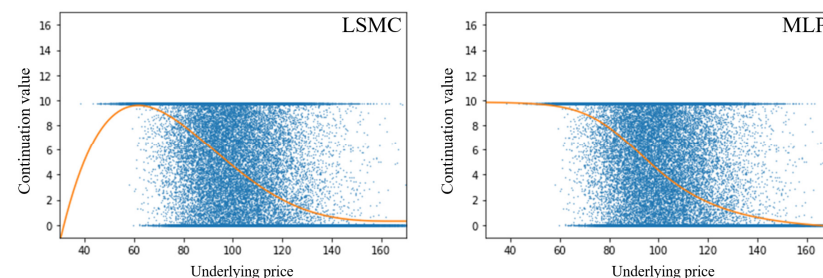


Figure 6. Continuation value functions (red lines) for the one-dimensional problem (with constant volatility) in the Bermudan capped option with exercisable dates $N = 2$. The blue dots represent the generated sample points of Bermudan option payoffs at expiration T_2 .

For the two-dimensional case with single-asset stochastic volatility for Bermudan capped put options, the continuation value functions have three-dimensional surfaces, as shown in Figure 7, wherein the left-hand and right-hand plots depict the continuation values with respect to volatility and underlying asset price directions for the LSMC method and the MLP, respectively. As in the one-dimensional problem, the payoff function is piecewise linear with respect to the underlying asset price direction and is flat when the underlying asset price exceeds the strike price $K = 105$ or is less than a certain value related to the capped rate $D = 10$. Since continuation value functions are supposed to approximate the payoff function at the maturity T_2 , given the information up to time T_1 , their surfaces are expected to have similar shapes, i.e., continuation values are approximately zero or close to the capped rate for larger or smaller values of the underlyings, respectively. In view of this payoff structure for the Bermudan capped put options, the continuation value

surface of the MLP seems to approximate the payoff function more accurately than that of the LSMC.

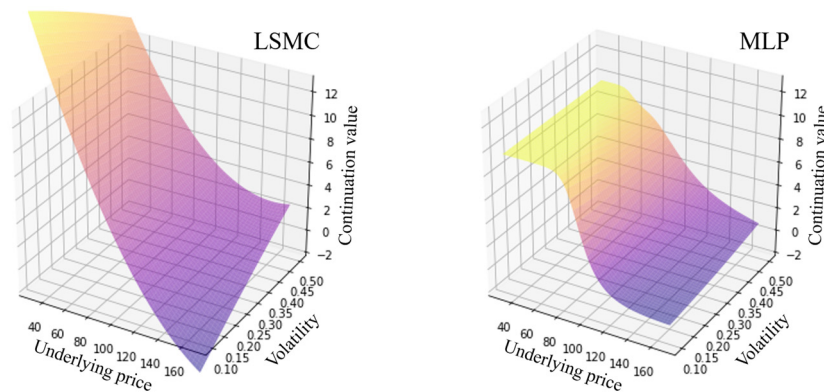


Figure 7. Continuation value surfaces at T_1 of the two-dimensional problem (single-asset stochastic volatility) for Bermudan capped put options.

Remark 1. In general, the visualization of nonparametric methods provides an intuitive interpretation of the estimated functions. We have observed that “the continuation value function of the MLP monotonically decreases with the underlying price, whereas the one obtained from the LSMC method is a non-monotonic function,” and that “since the payoff function of the Bermudan capped-put option is piecewise linear and it monotonically decreases with the underlying price, the monotonicity of the continuation value function is more consistent with the payoff structure of the Bermudan capped-put option,” as stated earlier in this section. Such a visualization helps in understanding the valuation structure for the applied method in the middle of the process for Bermudan option pricing, but the effect of the approximation error may be weakened in the total procedure. However, we should be able to understand the approximation errors of estimated continuation value functions intuitively in the middle of the process from such a visualization.

4.5. Comparison of Accuracy in Continuation Values

In the previous subsection, we observed that the continuation value surfaces of the MLP may approximate the payoff functions more accurately than those of the LSMC by visualizing the continuation value surfaces. To further investigate the estimated continuation value surfaces in higher-dimension problems, we next measure the accuracy of the continuation values using a four-dimensional problem of the Bermudan capped put basket option with two exercisable dates (i.e., $d_1 = 4$ and $N = 2$) for both the MLP and the LSMC method.

Consider a problem of estimating the continuation values at T_1 , as shown in Figure 8. Since Bermudan options with exercisable dates $N = 2$ become simple European options if not exercised at T_1 , the continuation values at T_1 may be estimated via European option prices expiring at T_2 , given the state variables at T_1 , i.e., $X_{T_1} = [S_{1,T_1}, \sigma_{1,T_1}, S_{2,T_1}, \sigma_{2,T_2}]^\top$. Therefore, we can calculate the accuracy of the estimated continuation values at T_1 by measuring the differences between the estimated continuation values and the European option prices at T_1 by specifying different state variables as input values of the European options.

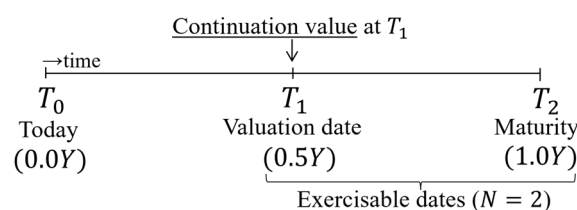


Figure 8. Continuation value at T_1 with two exercisable dates.

We first solved the Bermudan option problem with $d_t = 4$ and $N = 2$ using the same parameter settings as those in the previous subsections by applying the MLP and the LSMC methods; subsequently, we calculated the continuation values at T_1 , given the state variables specified in Table 8. Note that Table 8 defines the discretized domain of the state variables, wherein each state variable is discretized in the interval between the minimum and maximum values so that the number of grid points in 1 dimension becomes $2^4 + 1$. Then, the total number of grid points is given by 83,521 ($= 17^4$). Similarly, we applied the standard Monte Carlo simulation to compute the European option price on each grid and repeated this procedure 83,521 times to estimate the surface of the European option prices. This surface of the European option prices may be considered to provide an approximation of the theoretical continuation value (see notes in Table 8), and we can measure the accuracy of the continuation values using the difference between the estimated continuation values with the MLP and the LSMC methods and the simulation-based (theoretical) surface.

Table 8. State variables at T_1 . Each state variable is equally discretized into the number of points from minimum up to maximum value.

Variable	Minimum	Maximum	# of Points ¹	Interval
S_{1,T_1}	25.0	175.0	$2^4 + 1$	9.375
σ_{1,T_1}	0.05	0.55	$2^4 + 1$	0.03125
S_{2,T_1}	25.0	175.0	$2^4 + 1$	9.375
σ_{2,T_1}	0.05	0.55	$2^4 + 1$	0.03125

¹ The number of grid points with state variables at T_1 are 83,521 ($= 17^4$) in total. The number of sample paths generated for evaluating each European option in the standard Monte Carlo simulation is 10,000. The average values of estimated European prices and standard errors are 4.997 and 0.027, respectively, indicating that the 95% confidence interval is given by 4.997 ± 0.053 on average for the Monte Carlo simulations.

Additionally, we can change the number of hidden layers/neurons and the type of activation function in the MLP to verify their effects on its accuracy. In this study, we evaluate the size of accuracy in terms of the following normalized root-mean-squared error (NRMSE) for each methodology:

$$NRMSE = \frac{1}{\sqrt{I}} \frac{\sqrt{\sum_{i=1}^I (p_i - \hat{p}_i)^2}}{\hat{p}_{max} - \hat{p}_{min}}, \tag{11}$$

where I is the total number of grid points (i.e., $I = 83,521$), and p_i and \hat{p}_i are the i^{th} -continuation value and the corresponding European option price on the same grid point. In Equation (11), the root-mean-squared error is normalized by the difference between \hat{p}_{min} and \hat{p}_{max} , which are the minimum and maximum values of European option prices over the entire grid.

We computed the NRMSEs for different settings of neural networks in the case of the MLP, as shown in Table 9, wherein we changed the number of hidden layers/neurons and applied two types of activation functions, i.e., the ReLU and sigmoid functions. Note that the NRMSE with the LSMC method is also computed, as shown in the bottom row of the table, while Figure 9 compares the same NRMSE with respect to a different number of neurons for 16, 32, and 128 using bar graphs. In Table 9, we first observe that the MLP almost always provided better accuracy in terms of NRMSEs compared with the LSMC. Second, when comparing the types of activation functions, the MLP with the sigmoid function was always better than the MLP with the ReLU function for estimating continuation value surfaces when the number of hidden layers/neurons was fixed. This may be explained by the smoothness of the sigmoid function; the continuation value functions are expected to be smooth with respect to the state variables and can be fitted via smooth functions (e.g., the sigmoid function) better than non-smooth functions, such as the ReLU function. Third, an increase in the number of hidden layers is effective for a few hidden layers but does not necessarily improve the NRMSE when the number of hidden

layers is three or larger for both MLPs with the ReLU and sigmoid functions. However, in any case, we obtained better NRMSEs by using the MLP with the sigmoid function.

Table 9. NRMSE comparisons with continuation values at T_1 by network settings of the MLP. It differs by activation functions and the number of hidden layers/neurons.

		NRMSE									
# of Neurons		16		32		64		128		256	
# of Hidden Layers		ReLU	Sigmoid	ReLU	Sigmoid	ReLU	Sigmoid	ReLU	Sigmoid	ReLU	Sigmoid
1		0.266	0.075	0.262	0.099	0.198	0.160	0.244	0.162	0.160	0.211
2		0.176	0.101	0.145	0.078	0.161	0.081	0.154	0.073	0.311	0.100
3		0.215	0.083	0.246	0.072	0.126	0.097	0.157	0.083	0.271	0.085
4		0.128	0.115	0.225	0.098	0.172	0.079	0.124	0.059	0.184	0.082
5		0.137	0.097	0.132	0.100	0.169	0.090	0.132	0.127	0.154	0.076
LSMC						0.244					

note NRMSE = normalized root-mean-squared error.

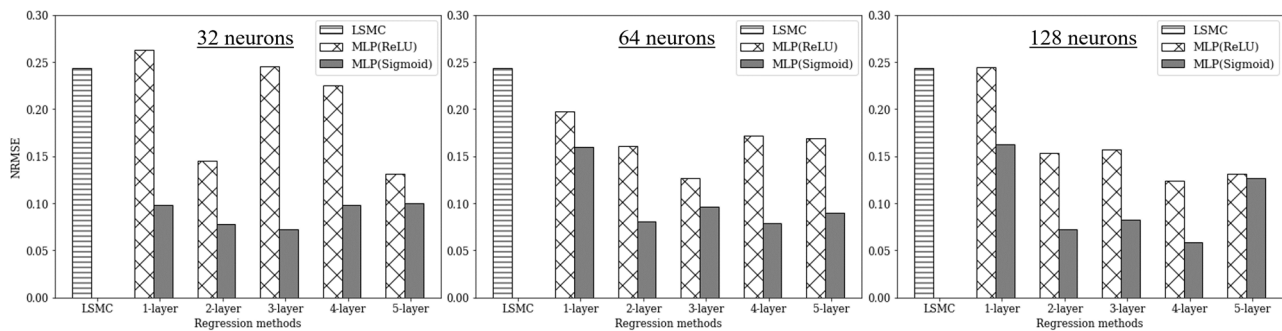


Figure 9. NRMSE comparison with continuation values at T_1 . MLP differences in a number of hidden layers, neurons, and selected activation functions.

5. Discussion of Robustness and Computational Costs

In this section, we discuss the robustness and computational costs of our experiment for the pricing algorithm of multi-asset Bermudan commodity options using the MLP.

First, we discuss the learning adaptability of the MLP applied to our multi-asset Bermudan commodity option problems with stochastic volatility. Figure 10 depicts changes in the mean and standard deviation of learning rates with respect to the number of epochs in the MLP. In this figure, we see that training and validation losses remain close, which indicates no overfitting. Furthermore, the learning rate decreases rapidly until the number of epochs is 10 and stays at sufficiently good levels thereafter.

Next, we estimated the computational costs of the learning and valuation procedure when the problem dimension is increased. Table 10 compares the computational costs with respect to the dimensions of $d_l = 2, 4, 8, 16, 32, 64$, wherein the same numerical experiment as that of the previous section was repeated 100 times and computed the mean and standard deviation of the computational time for each algorithm. Furthermore, the average numbers of epochs in learning are also computed in the MLP. Although the LSMC generally performs much better in terms of computational costs when the dimension is particularly low, its computational time grows exponentially in tandem with the size of the dimension. This is because when using a polynomial regression in the LSMC, the number of terms in the polynomial function increases combinatorially with the number of variables, even though its maximum order is fixed.

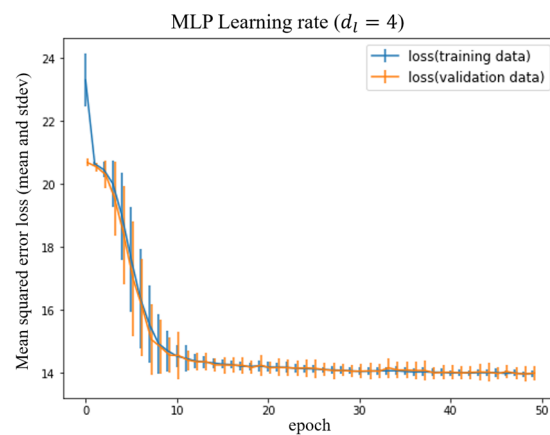


Figure 10. Learning rate by epochs. The learning of the MLP was repeated 100 times and computed the mean (line) and standard deviation (bar) of learning losses in the Bermudan capped put option with the exercise dates $N = 2$ and two underlying assets with stochastic volatility (i.e., $d_n = 4$).

Table 10. Computational cost comparisons. Each algorithm’s computational time periods (seconds) and their statistics by dimensions were calculated 100 times in the same option condition, as shown in Figure 10. In the MLP, the average numbers of epochs in learning are also listed.

Dimensions (d_l)	LSMC [Sec.]				MLP [Sec.] ²				
	Learning ¹		Valuation		Learning		Valuation		
	Mean	(St. Dev.)	Mean	(St. Dev.)	Mean	(St. Dev.)	# of Epochs	Mean	(St. Dev.)
2	0.025	(0.002)	0.003	(0.003)	20.038	(3.599)	186.19	0.006	(0.002)
4	0.076	(0.019)	0.003	(0.001)	17.637	(4.403)	156.56	0.006	(0.002)
8	0.257	(0.064)	0.005	(0.001)	17.382	(5.089)	157.43	0.006	(0.001)
16	0.861	(0.083)	0.020	(0.004)	9.773	(7.707)	81.99	0.007	(0.002)
32	3.708	(0.334)	0.068	(0.020)	5.114	(2.829)	38.53	0.008	(0.002)
64	19.887	(2.842)	0.210	(0.018)	5.367	(1.468)	36.58	0.009	(0.002)

¹ LSMC with a multi-dimensional quadratic function. ² MLP with 2 hidden layers with 64 neurons of sigmoid activation functions.

In contrast with the LSMC, the computational cost in the MLP is mostly unaffected by the size of the dimension but is directly proportional to the average number of epochs in learning. The computational cost of learning depends on how often the networks are updated during training, but the computation cost per one cycle of training data (i.e., epoch) remains the same when the size of the network is fixed. In the numerical experiments, we applied two hidden layers with sixty-four neurons using sigmoid activation functions, whereby the computational cost per epoch remained almost the same regardless of dimensions; the computational time in learning is determined by the total number of epochs. Although the computational cost per epoch slightly increases as the number of features in the input layer increases by dimension, the average computational time decreases even for large dimensions with the reduction in the average number of epochs due to the early stopping rule. This is the benefit of introducing the early stopping rule in Algorithm 1, which is particularly effective for higher-dimensional problems to avoid unnecessarily increasing training iterations (and overfitting). Note that the average computational time for both learning and valuation became smaller for the MLP than the LSMC when $d_l = 64$.

6. Conclusions

In this study, we detailed the use of a neural network for pricing multi-asset Bermudan option problems with stochastic volatility in commodity markets and illustrated its effectiveness. First, we employed the MLP to estimate continuation values in the multi-dimensional Bermudan commodity option problem, whereby we formulated the multi-asset stochastic

volatility model and generated Monte Carlo simulation sample paths for learning continuation value functions using the MLP. Then, in the applied algorithm, we introduced early stopping into the learning of the MLP to avoid unnecessarily increasing training iterations and overfitting. The early stopping rule was activated by counting the number of epochs without improvement compared with the previous epoch. We also introduced a resampling process, and a valuation procedure was applied for the estimated neural networks by generating a different set of simulation sample paths. We executed numerical experiments to evaluate the accuracy of the continuation values and the initial price of the option using different settings of networks, problem dimension, and exercisable dates, whereby two types of payoff functions for Bermudan commodity options were considered, namely Bermudan put options and Bermudan capped put options. From our numerical analysis, we clarified the following observations:

1. No significant difference was observed between the MLP and the standard LSMC method when solving Bermudan put option problems with a few exercisable dates. However, there was a slight difference for the Bermudan capped put options; this difference was emphasized when the number of exercisable dates increased. A similar tendency was observed for higher-dimensional cases, but the gap narrowed between the mean values of the two methods for Bermudan capped put options, as shown in our additional numerical experiments.
2. While it turned out that the MLP was not much better than the standard LSMC from the numerical experiments in Section 4.3 for high-dimensional cases, it is meaningful to show how the accuracy and computational time can be achieved using the current computational resources. In addition, we expect that the MLP has the potential to achieve much better accuracy due to its generality and flexibility. Moreover, if computational power is increased, the MLP should become more efficient since computational effort grows slower than that of polynomial regressions in the standard LSMC for higher-dimensional problems, as illustrated in the numerical experiments in Section 5.
3. From the perspective that the continuation value function is expected to approximate the payoff function (given state variables) one step before maturity, the shape of the continuation values from the MLP reflected the structure of payoff functions more accurately than the LSMC method.
4. Based on the fact that the continuation values of Bermudan options one step before maturity can be computed as European option prices, we measured the accuracy of the estimated continuation values and examined the effects of different network configurations in the MLP, changing the number of hidden layers/neurons and the choice of activation functions. We observed that the MLP almost always provided better accuracy in terms of NRMSEs compared with the LSMC; furthermore, when comparing the types of activation functions, the MLP with the sigmoid function was always better than the MLP with the ReLU function for estimating continuation value surfaces. An increase in the number of hidden layers was effective for a few network layers but did not necessarily improve the accuracy when the number of hidden layers was three or larger.
5. We computed the learning rate by epochs to show the learning adaptability of our proposed algorithm using the MLP, which indicated no overfitting and achieved sufficiently good levels of learning rates at approximately 10 epochs. Additionally, we showed that although the LSMC generally performs significantly better in terms of computational costs when the dimension is particularly low, its computational time grows exponentially with the size of the dimension due to the combinatorial characterization with respect to the number of terms in the polynomial functions. Conversely, the computational costs of the MLP were mostly unaffected by the size of the dimension or even decreased for large dimensions due to the introduction of the early stopping rule.

Essentially, the use of neural networks for option pricing has the advantage of recognizing sizable input features and generating flexible output features in a unified framework. Nevertheless, there are some drawbacks: high computational effort and resources are required for learning networks, especially for exotic options, including the Bermudan commodity options considered in this study. Additionally, it is necessary to frequently re-learn the networks in response to market conditions. However, we observed that the neural network approach using the MLP reached an appropriate level of learning rates at around 10 epochs, even in high-dimensional cases, as illustrated in our numerical experiments. Therefore, this approach is expected to reduce learning costs if network configurations are developed appropriately.

Although this study chose a relatively simple structure of multi-layered networks, there are other types of network structures such as a recursive structure and unsupervised learning, as discussed in various fields, including pattern recognition and time-series prediction. In finance, although some examples of recursive neural networks for time-series analyses exist, to the best of our knowledge, their use for option pricing has not been considered sufficiently. Moreover, it is important to use empirical data and demonstrate the practicability and applications for risk management in actual commodity market businesses. Such further investigation is interesting and could be considered potential topics for future study.

Author Contributions: Conceptualization, K.H. and Y.Y.; methodology, K.H. and Y.Y.; software, K.H.; validation, K.H. and Y.Y.; formal analysis, K.H.; investigation, K.H.; resources, K.H. and Y.Y.; data curation, K.H.; writing—original draft preparation, K.H.; writing—review and editing, Y.Y.; visualization, K.H.; supervision, Y.Y.; project administration, Y.Y.; funding acquisition, Y.Y. All authors have read and agreed to the published version of the manuscript.

Funding: Grant-in-Aid for Scientific Research (A) 20H00285 and Grant-in-Aid for Challenging Research (Exploratory) 19K22024 from the Japan Society for the Promotion of Science (JSPS).

Data Availability Statement: Not applicable.

Acknowledgments: This work was supported by the Grant-in-Aid for Scientific Research (A) 20H00285 and Grant-in-Aid for Challenging Research (Exploratory) 19K22024 from the Japan Society for the Promotion of Science (JSPS).

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A. Recursive Formulation for a Bermudan Option Pricing Problem

Given a complete filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{0 \leq t \leq T}, \mathbb{P})$ with a finite time horizon $T > 0$, we assume that a set of underlying assets is modeled using a multifactor process $(X_t)_{0 \leq t \leq T}$ adapted to the filtration, $(\mathcal{F}_t)_{0 \leq t \leq T}$, and that \mathbb{P} is an associated risk-neutral measure. We consider a Bermudan option with exercise dates $0 = T_0 \leq T_1 < T_2 < \dots < T_N = T$ and discrete-time payoff process P_{T_n} if exercised at times $(T_n)_{0 \leq n \leq N}$, where P_{T_n} is specified as a function of X_{T_n} .

In the Bermudan option, the continuation and exercise values are compared at each exercisable period, while the option is exercised if the exercise value is higher. Therefore, Bermudan option value V_{T_n} is computed using the following recursive equation:

$$\begin{cases} V_{T_N} = P_{T_N} \\ V_{T_{n-1}} = \max(P_{T_{n-1}}, e^{-r\delta_{T_n}} \mathbb{E}[V_{T_n} | \mathcal{F}_{T_{n-1}}]) \end{cases}, 1 \leq n \leq N \tag{A1}$$

In the risk-neutral measure, the conditional expected value of the risk-neutral probability measure $\tilde{\mathbb{P}}$ with the risk-free interest rate r and the interval between T_{n-1} and T_n as δ_{T_n} , based on the filtration $\mathfrak{F}_{T_{n-1}}$ up to time T_{n-1} , indicates the continuation value $U_{T_{n-1}}$ as

$$U_{T_{n-1}} = e^{-r\delta_{T_n}} \tilde{\mathbb{E}}[V_{T_n} | \mathfrak{F}_{T_{n-1}}]. \tag{A2}$$

Bermudan option value $V_{T_{n-1}}$ at time T_{n-1} is sequentially calculated backward when the continuation value $U_{T_{n-1}}$ is identified. The key is estimating the continuation value as a function that consists of underlying multivariate risk factors. Since underlying assets consist of a multi-dimensional Markov process, the continuation value function can be expressed as the multi-dimensional nonlinear function with the Markov process state variables $X_{T_{n-1}}$,

$$\tilde{E}[V_{T_n} | \mathfrak{F}_{T_{n-1}}] = \tilde{E}[V_{T_n} | X_{T_{n-1}}]. \tag{A3}$$

Furthermore, from the definition of the conditional expectation, there is a measurable function $h_{T_{n-1}}$ that satisfies the following equation:

$$h_{T_{n-1}}(X_{T_{n-1}}) = e^{-r\delta T_n} \tilde{E}[V_{T_n} | X_{T_{n-1}}]. \tag{A4}$$

For the approximation of a function $h_{T_{n-1}}(X_{T_{n-1}})$, we can consider $\Phi_{T_{n-1}}(x_{T_{n-1}})$ as the approximation function at time T_{n-1} ,

$$h_{T_{n-1}}(X_{T_{n-1}}) \approx \Phi_{T_{n-1}}(X_{T_{n-1}}). \tag{A5}$$

After that, the price at time $t = 0$ is calculated by following recursive backward procedures using the relationship V_{T_n} and $V_{T_{n-1}}$. At maturity $T_N (= T)$, the continuation value of the Bermudan option is $U_T \equiv 0$. Therefore, the Bermudan option's value at maturity T_N is:

$$V_{T_N} = P_{T_N}. \tag{A6}$$

By (A4), the continuation value $U_{T_{N-1}}$ at time T_{N-1} is expressed as

$$U_{T_{N-1}} = h_{T_{N-1}}(X_{T_{N-1}}) \approx \Phi_{T_{N-1}}(X_{T_{N-1}}). \tag{A7}$$

The Bermudan option value at time T_{N-1} is expressed as

$$V_{T_{N-1}} := \max(P_{T_{N-1}}, \Phi_{T_{N-1}}(X_{T_{N-1}})). \tag{A8}$$

We can obtain Bermudan option value V_{T_0} by adapting (A7) and (A8) backward, recursively, each time step to $n = 1$. In multi-asset Bermudan option pricing, Monte Carlo simulations are generally used because other numerical methods become exponentially more difficult in higher-dimensional cases. By simulating a large number of paths, we can use the average of the prices obtained from each path as an estimator of the price as

$$\overline{V}_{T_0} = \frac{1}{M} \sum_{j=1}^M V_{T_0}^{(j)}, \tag{A9}$$

where M is the number of simulated paths.

From the above, the prices of the Bermudan options can be obtained by finding the approximate functions of the continuation value functions at each exercisable period.

Appendix B. Least-Squares Monte Carlo Method

The least-squares Monte Carlo (LSMC) method, proposed by Longstaff and Schwartz (2001), is a method of early-exercisable option pricing in which regression calculation uses simulation sample paths. In the LSMC method, a polynomial function of the Markov process state variables is applied to identify the continuation values. The following is an algorithm for Bermudan option pricing using the LSMC method.

Step 0. Generate Monte Carlo sample paths of the underlying asset prices and state variables. We denote the underlying asset prices at time t in the j -th sample path as $S_t^{(j)}$

and the Markov process state variables as $x_t^{(j)}$. Subsequently, we obtain the series of paths as

$$\begin{pmatrix} S_{T_1}^{(1)}, & S_{T_2}^{(1)}, & \dots, & S_{T_N}^{(1)} & x_{T_1}^{(1)}, & x_{T_2}^{(1)}, & \dots, & x_{T_N}^{(1)} \\ S_{T_1}^{(2)}, & S_{T_2}^{(2)}, & \dots, & S_{T_N}^{(2)} & x_{T_1}^{(2)}, & x_{T_2}^{(2)}, & \dots, & x_{T_N}^{(2)} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ S_{T_1}^{(M)}, & S_{T_2}^{(M)}, & \dots, & S_{T_N}^{(M)} & x_{T_1}^{(M)}, & x_{T_2}^{(M)}, & \dots, & x_{T_N}^{(M)} \end{pmatrix}. \tag{A10}$$

Step 1. Calculate the series of Bermudan option values at maturity $T_N (= T)$, as follows:

$$V_{T_N} := \left[g\left(S_{T_N}^{(1)}\right), g\left(S_{T_N}^{(2)}\right), \dots, g\left(S_{T_N}^{(M)}\right) \right]^T, \tag{A11}$$

where g denotes a payoff function of S_t .

Step 2. Find a polynomial function that approximates the continuation values. Herein, we denote a polynomial function as $\hat{h}_{T_{N-1}}$ and a measurable function in (A4) as $h_{T_{N-1}}$. Then,

$$h_{T_{N-1}}(x_{T_{N-1}}) \approx \hat{h}_{T_{N-1}}(x_{T_{N-1}}). \tag{A12}$$

Additionally, $\hat{h}_{T_{N-1}}$ is sought to minimize the following equation:

$$\frac{1}{M} \sum_{i=1}^M \left(\hat{h}_{T_{N-1}}(x_{T_{N-1}}^{(i)}) - e^{-r\delta_{T_N}} V_{T_N} \right)^2. \tag{A13}$$

Step 3. Calculate the approximated continuation values. Let $h_{T_{N-1}} \equiv \hat{h}_{T_{N-1}}$ and set the series of approximated continuation values at T_{N-1} as

$$\left[\hat{h}_{T_{N-1}}\left(x_{T_{N-1}}^{(1)}\right), \hat{h}_{T_{N-1}}\left(x_{T_{N-1}}^{(2)}\right), \dots, \dots, \hat{h}_{T_{N-1}}\left(x_{T_{N-1}}^{(M)}\right) \right]^T. \tag{A14}$$

Step 4. Calculate the exercised values and Bermudan option values. The series of exercise values at T_{N-1} using underlying asset prices (A10) is

$$\left[g\left(S_{T_{N-1}}^{(1)}\right), g\left(S_{T_{N-1}}^{(2)}\right), \dots, g\left(S_{T_{N-1}}^{(M)}\right) \right]^T. \tag{A15}$$

Then, the series of Bermudan option values $V_{T_{N-1}}$ at time T_{N-1} is obtained as

$$\left[V_{T_{N-1}}\left(S_{T_{N-1}}^{(1)}, x_{T_{N-1}}^{(1)}\right), V_{T_{N-1}}\left(S_{T_{N-1}}^{(2)}, x_{T_{N-1}}^{(2)}\right), \dots, \dots, V_{T_{N-1}}\left(S_{T_{N-1}}^{(M)}, x_{T_{N-1}}^{(M)}\right) \right]^T, \tag{A16}$$

where

$$V_{T_{N-1}}\left(S_{T_{N-1}}^{(j)}, x_{T_{N-1}}^{(j)}\right) := \max\left(g\left(S_{T_{N-1}}^{(j)}\right), h_{T_{N-1}}\left(x_{T_{N-1}}^{(j)}\right)\right), j = 1, \dots, M, \tag{A17}$$

Step 5. Repeat Step 2 to Step 4 for possible exercise times T_{N-1}, T_{N-2}, \dots , until time T_0 . A series of Bermudan option values V_{T_0} at time T_0 are obtained by repeating Step 2 to Step 4 backward to time T_0 .

Step 6. Calculate the Bermudan option price \overline{V}_{T_0} . Equation (A9) gives the Bermudan option price \overline{V}_{T_0} at time T_0 .

Using this approach, we calculated the Bermudan option price using the continuation values from the polynomial function at each exercisable time. The main point of the pricing procedure in the LSMC method is that the polynomial function is defined to approximate the continuation values at each exercisable time point.

References

- Abadi, Martín, Ashish Agarwal, Paul Barham, Eugene Brevdo, Zhiheng Chen, Craig Citro, Greg S. Corrado, Andy Davis, Jeffrey Dean, Matthieu Devin, and et al. 2015. TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems. TensorFlow. Available online: <https://www.tensorflow.org/> (accessed on 1 December 2022).
- Andersson, Kristoffer, and Cornelis W. Oosterlee. 2021. A Deep Learning Approach for Computations of Exposure Profiles for High-Dimensional Bermudan Options. *Applied Mathematics and Computation* 408: 126332. [CrossRef]
- Arnold, Vladimir I. 2009. On functions of three variables. Collected Works: Representations of Functions. *Celestial Mechanics and KAM Theory* 1957–1965: 5–8.
- Ball, Clifford A., and Antonio Roma. 1994. Stochastic volatility option pricing. *Journal of Financial and Quantitative Analysis* 29: 589–607. [CrossRef]
- Ballestra, Luca Vincenzo, and Graziella Pacelli. 2013. Pricing European and American Options with Two Stochastic Factors: A Highly Efficient Radial Basis Function Approach. *Journal of Economic Dynamics and Control* 37: 1142–67. [CrossRef]
- Becker, Sebastian, Patrick Cheridito, and Arnulf Jentzen. 2020. Pricing and Hedging American-Style Options with Deep Learning. *Journal of Risk and Financial Management* 13: 158. [CrossRef]
- Becker, Sebastian, Patrick Cheridito, Arnulf Jentzen, and Timo Welti. 2021. Solving High-Dimensional Optimal Stopping Problems Using Deep Learning. *European Journal of Applied Mathematics* 32: 470–514. [CrossRef]
- Chen, Yangang, and Justin W. L. Wan. 2021. Deep Neural Network Framework Based on Backward Stochastic Differential Equations for Pricing and Hedging American Options in High Dimensions. *Quantitative Finance* 21: 45–67. [CrossRef]
- Choon, Ong Hong, Leong Chee Hoong, and Tai Sheue Huey. 2008. A functional approximation comparison between neural networks and polynomial regression. *WSEAS Trans. Math* 7: 353–63.
- Clark, Iain J. 2014. *Commodity Option Pricing: A Practitioner's Guide*. New York: John Wiley & Sons.
- Cybenko, George. 1989. Approximation by Superpositions of a Sigmoidal Function. *Mathematics of Control, Signals and Systems* 2: 303–14. [CrossRef]
- Feng, Guiyun, Guangwu Liu, and Lihua Sun. 2013. A Nonparametric Method for Pricing and Hedging American Options. In *2013 Winter Simulations Conference (WSC)*. Piscataway: IEEE, pp. 691–700.
- Google. 2022. Frequently Asked Questions. Available online: <https://research.google.com/colaboratory/faq.html> (accessed on 1 December 2022).
- Hahn, Warren J., and James S. Dyer. 2008. Discrete time modeling of mean-reverting stochastic processes for real option valuation. *European Journal of Operational Research* 184: 534–48. [CrossRef]
- Haugh, Martin B., and Leonid Kogan. 2004. Pricing American Options: A Duality Approach. *Operations Research* 52: 258–70. [CrossRef]
- Heston, Steven L. 1993. A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options. *The Review of Financial Studies* 6: 327–43. [CrossRef]
- Kingma, Diederik P., and Jimmy Ba. 2014. Adam: A Method for Stochastic Optimization. *arXiv* 1412: 6980.
- Kohler, Michael, Adam Krzyżak, and Nebojsa Todorovic. 2010. Pricing of High-Dimensional American Options by Neural Networks. *Mathematical Finance* 20: 383–410. [CrossRef]
- Kolmogorov, Andrei Nikolaevich. 1957. On the Representation of Continuous Functions of Many Variables by Superposition of Continuous Functions of One Variable and Addition. *Doklady Akademii Nauk* 114: 953–56.
- Lapeyre, Bernard, and Jérôme Lelong. 2021. Neural Network Regression for Bermudan Option Pricing. *Monte Carlo Methods and Applications* 27: 227–47. [CrossRef]
- Li, Nan. 2022. An Iteration Algorithm for American Options Pricing based on Reinforcement Learning. *Symmetry* 14: 1324. [CrossRef]
- Liang, Jian, Zhe Xu, and Peter Li. 2021. Deep Learning-Based Least Squares Forward-Backward Stochastic Differential Equation Solver for High-Dimensional Derivative Pricing. *Quantitative Finance* 21: 1309–23. [CrossRef]
- Lin, Jingying, and Caio Almeida. 2021. American Option Pricing with Machine Learning: An Extension of the Longstaff-Schwartz Method. *Brazilian Review of Finance* 19: 85–109. [CrossRef]
- Liu, Shuaiqiang, Cornelis W. Oosterlee, and Sander M. Bohte. 2019. Pricing Options and Computing Implied Volatilities Using Neural Networks. *Risks* 7: 16. [CrossRef]
- Longstaff, Francis A., and Eduardo S. Schwartz. 2001. Valuing American Options by Simulation: A Simple Least-Squares Approach. *The Review of Financial Studies* 14: 113–47. [CrossRef]
- Ludkovski, Michael. 2018. Kriging Metamodels and Experimental Design for Bermudan Option Pricing. *Journal of Computational Finance* 22: 37–77. [CrossRef]
- Montesdeoca, Luis, and Mahesan Niranjana. 2016. Extending the Feature Set of a Data-Driven Artificial Neural Network Model of Pricing Financial Options. Paper presented at 2016 IEEE Symposium Series on Computational Intelligence (SSCI), Athens, Greece, December 6–9; pp. 1–6.
- Pedregosa, Fabian, Gaël Varoquaux, Alexandre Gramfort, Vincent Michel, Bertrand Thirion, Olivier Grisel, Mathieu Blondel, Peter Prettenhofer, Ron Weiss, Vincent Dubourg, and et al. 2011. Scikit-Learn: Machine Learning in Python. *The Journal of Machine Learning Research* 12: 2825–30.
- Puka, Radosław, Bartosz Łamasz, and Marek Michalski. 2021. Effectiveness of Artificial Neural Networks in Hedging Against WTI Crude Oil Price Risk. *Energies* 14: 3308. [CrossRef]

- Ruf, Johannes, and Weiguan Wang. 2020. Neural Networks for Option Pricing and Hedging: A Literature Review. *Journal of Computational Finance* 24: 1–46. [CrossRef]
- Salvador, Beatriz, Cornelis W. Oosterlee, and Remco van der Meer. 2020. Financial Option Valuation by Unsupervised Learning with Artificial Neural Networks. *Mathematics* 9: 46. [CrossRef]
- Schwartz, Eduardo S. 1997. The Stochastic Behavior of Commodity Prices: Implications for Valuation and Hedging. *The Journal of Finance* 52: 923–73. [CrossRef]
- Serpen, Gursel, and Zhenning Gao. 2014. Complexity analysis of multilayer perceptron neural network embedded into a wireless sensor network. *Procedia Computer Science* 36: 192–97. [CrossRef]
- Sirignano, Justin, and Konstantinos Spiliopoulos. 2018. Dgm: A Deep Learning Algorithm for Solving Partial Differential Equations. *Journal of Computational Physics* 375: 1339–64. [CrossRef]
- Tavella, Domingo, and Curt Randall. 2000. *Pricing Financial Instruments: The Finite Difference Method*. New York: John Wiley & Sons.
- Yamada, Yuji. 2012. Properties of Optimal Smooth Functions in Additive Models for Hedging Multivariate Derivatives. *Asia-Pacific Financial Markets* 19: 149–79. [CrossRef]
- Yamada, Yuji. 2017. Optimal Hedging of Basket Barrier Options with Additive Models and Its Application to Equity Value Separation Problem. *Asia-Pacific Financial Markets* 24: 1–18. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

MDPI
St. Alban-Anlage 66
4052 Basel
Switzerland
www.mdpi.com

Journal of Risk and Financial Management Editorial Office

E-mail: jrfm@mdpi.com
www.mdpi.com/journal/jrfm



Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.



Academic Open
Access Publishing

mdpi.com

ISBN 978-3-0365-9028-8